AFRL-RI-RS-TR-2015-204

MULTIAGENT TASK COORDINATION USING A DISTRIBUTED OPTIMIZATION APPROACH

BETHUNE-COOKMAN UNIVERSITY

SEPTEMBER 2015

FINAL TECHNICAL REPORT

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

STINFO COPY

AIR FORCE RESEARCH LABORATORY
INFORMATION DIRECTORATE

■ **AIR FORCE MATERIEL COMMAND** ■ **UNITED STATES AIR FORCE** ■ **ROME, NY 13441**

NOTICE AND SIGNATURE PAGE

AFRL-RI-RS-TR-2015-204 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

FOR THE DIRECTOR:

/ S /
GENNADY STASKEVICH
Work Unit Manager

/ S /
JULIE BRICHACEK, Chief
Information Systems Division
Information Directorate

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
SEPTEMBER 2015	FINAL TECHNICAL REPORT	MAR 2013 – MAR 2015

4. TITLE AND SUBTITLE

MULTIAGENT TASK COORDINATION USING A DISTRIBUTED OPTIMIZATION APPROACH

5a. CONTRACT NUMBER FA8750-13-1-0109

5b. GRANT NUMBER N/A

5c. PROGRAM ELEMENT NUMBER 62788F

6. AUTHOR(S)

Jing Wang, Morrison Obeng, and Thomas Yang

5d. PROJECT NUMBER S2MA

5e. TASK NUMBER SA

5f. WORK UNIT NUMBER BC

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Bethune-Cookman University Inc.
Sponsored Programs
640 Dr Mary McLeod Bethune Blvd.
Daytona Beach, FL 32114-3012

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Air Force Research Laboratory/RISC
525 Brooks Road
Rome NY 13441-4505

10. SPONSOR/MONITOR'S ACRONYM(S) AFRL/RI

11. SPONSOR/MONITOR'S REPORT NUMBER AFRL-RI-RS-TR-2015-204

12. DISTRIBUTION AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited. PA# 88ABW-2015-3918
Date Cleared: 6 AUG 15

13. SUPPLEMENTARY NOTES

14. ABSTRACT

In this project, several fundamental research topics have been carried out for developing multiagent task coordination strategies under a distributed optimization framework. The proposed subjects are critical to the development of engineered multiagent systems such as robotic networks, sensor networks, and computer networks, and they are important to both military and civilian applications. The objectives of the proposed research are three-folds: perform systematic controllability analysis for multiagent networks which may have nonlinear dynamics, design distributed optimal and adaptive coordination protocols in the presence of various model and communication uncertainties, and conduct computer simulation and experimental validation of the proposed designs using mobile robotic platforms. The project renders novel methods for discontinuous cooperative control under least restrictive sensing/communications and communication delays, approximate dynamic programming based optimal cooperative control, adaptive cooperative control of uncertain multiagent systems, and distributed formation control and coverage controls of multiple mobile robots with kinematic constraints.

15. SUBJECT TERMS

Multiagent coordination, optimization, distributed control, learning

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**GENNADY STASKEVICH**
U	U	U	SAR	90	**19b. TELEPHONE NUMBER** (*Include area code*) **315-330-4889**

TABLE OF CONTENTS

List of Figures

1.0 SUMMARY

In this project, several fundamental research topics have been carried out for developing multiagent task coordination strategies under a distributed optimization framework. The proposed subjects are critical to the development of engineered multiagent systems such as robotic networks, sensor networks, and computer networks, and they are important to both military and civilian applications. The objectives of the proposed research are three-folds: perform systematic controllability analysis for multiagent networks which may have nonlinear dynamics, design distributed optimal and adaptive coordination protocols in the presence of various model and communication uncertainties, and conduct computer simulation and experimental validation of the proposed designs using mobile robotic platforms.

The project renders novel methods for discontinuous cooperative control under least restrictive sensing/communications and communication delays, approximate dynamic programming based optimal cooperative control, adaptive cooperative control of uncertain multiagent systems, and distributed formation control and coverage controls of multiple mobile robots with kinematic constraints. In particular, the following three sets of results have been obtained:

- By analyzing the least restrictive condition for sensing/communication among multiagents, it is revealed that network consensus may not be achieved in the presence of discontinuous system dynamics. To address the issue, a new discontinuous cooperative law was proposed to achieve the task coordination of multiagent systems under directed and switching sensing/communication topologies [26]. In particular, we have shown the resilience of the proposed nominal cooperative control to certain extent in terms of communication link failures and time-delays.

- The optimal and adaptive task coordination for multiagent systems was thoroughly studied from three aspects. First, a practically implementable valued function approximation-based multiagent policy iteration algorithm was proposed for the optimal cooperative control of a class of nonlinear systems [31]. In the design, system behaviors are quantized using individual cost functions in order to direct the optimal operation of multiagent systems. Second, to further relax condition for the requirement of system dynamics, approximate Q function-based multiagent coordination algorithm was proposed. Third, the adaptive coordination of multiagent systems with uncertainties was studied. A new distributed adaptive cooperative control was proposed to deal with model uncertainties using neural network approximation and adaptive estimation of unknown parameters.

- Simulation and experimental study was conducted to test the robustness of the proposed coordination controls for multiagents based on case studies for formation control and coverage control of multiple mobile robots [27, 8, 30]. Specifically, for formation control of multiple mobile robots, both linearization-based design and nonlinear model-based design were proposed by assuming that only limited information of a desired trajectory is available. For coverage control, we proposed a distributed deployment algorithm for mobile robots to cover a convex region. The proposed deployment algorithm iteratively updates the Voronoi

partition through local information exchange, and then moves toward its centroid based on centroid-driven control algorithms.

Technical discussions of these three topics and research results are provided in the following sections.

2.0 INTRODUCTION

Multiagent systems are generically defined as a group of dynamical systems in which certain emergent behaviors are exhibited through local interactions of group members that individually have the capability of self-operating [18][13][2][19]. The fundamental issues in the study of multiagent systems are the analysis of network controllability and the design of coordination control protocol in order to achieve autonomous and optimal tasking allocation. For network controllability, the objective is to figure out the connectivity conditions on sensor/communication topologies among agents (including human operators) to achieve desired behaviors. Recent results on connectivity conditions for multiagent systems mostly assume perfect communication conditions, while in practice there often exist communication uncertainties and bandwidth limitations. For coordination control protocol design, the objective is to develop the proper control protocol to perform the coordinated tasks. The existing results in literature may not be directly applicable to multiagent task allocation due to possible link errors, long communication delays, and system uncertainties. More importantly, considering the possible dynamic task evolution for multiagent networks, the individual agent may exhibit multi-modal dynamics under different running circumstances or due to uncertainties and disturbances. All of these pose challenges in the design of performance guaranteed distributed coordination protocols that explicitly take into consideration system dynamics and uncertainties.

The proposed research have been centered on addressing the aforementioned fundamental issues by targeting the following objectives: perform systematic controllability analysis for multiagent networks which may have nonlinear dynamics, design distributed optimal and adaptive coordination protocols in the presence of various model and communication uncertainties, and conduct computer simulation and experimental validation of the proposed designs using mobile robotic platforms.

The project has rendered several significant results in the development of distributed optimal, adaptive and robust cooperative control protocols for the task coordination of multiagent systems. These results have reported and published in a number of IEEE conferences [26, 31, 27, 8, 30], and journal versions are under preparation for submission. The overall contributions of this project lie in two aspects: 1) the learning approaches borrowed from rich results in artificial intelligence research are effectively integrated with the rigorous control systems analysis tools, and produced novel approximate dynamic programming based optimal cooperative control and adaptive cooperative control for uncertain multiagent networks; 2) the research outcome on new task coordination algorithms for multiple agents operating in complex environments are a manifestation of robust intelligence.

The rest of the report is organized as follows. Section 3.0 presents the basic methods, assumptions and procedures in this research, and formulates the general multiagent coordination problem. Sections 4.0, 5.0, and 6.0 present the technical results and discussions on three sets of results in term of optimal and adaptive cooperative control design and applications, respectively. In each section, simulation and experimental results are given to illustrate the effectiveness of the proposed designs. Section 7.0 concludes the report.

3.0 METHODS, ASSUMPTIONS, AND PROCEDURES

3.1 Multiagent Dynamics and Assumptions

This research builds upon rigorous methods ranging from systems and controls theory, distributed reinforcement learning, adaptive learning to neural network approximation. We consider a set of agents $Q = \{1, \cdots, N\}$, where N is the number of agents in the group and assume that each agent evolves according to the general system dynamics described by

$$\left\{ \begin{array}{lcl} \dot{\xi}_i(t) & = & f_i(\xi_i, u_i, t) + \Delta f_i(\xi_i, t) + w_i(t), \\ y_i(t) & = & h_i(\xi_i, \xi_j) + v_i(t) \end{array} \right. \tag{1}$$

where $i = 1, \cdots, N$, $\xi_i(t) \in \Re^{n_i}$ is the state vector, $u_i \in \Re^{m_i}$ is the input vector, $m_i < n_i$, $y_i \in \Re^{p_i}$ is the output (measurement) vector, $w_i(t) \in \Re^{n_i}$ and $v_i(t) \in \Re^{p_i}$ are Gaussian noises with zero mean, $f_i(\cdot)$ and $h_i(\cdot)$ are piecewise continuous vector-valued functions of ξ_i on $\Re^{n_i}$, and $\Delta f_i(\cdot)$ and $\Delta g_i(\cdot)$ denote model uncertainties. Agent dynamics considered in (1) are given by the first-order differential equations in continuous time. The analog of (1) in discrete time can be defined by a system of first-order discrete time equations of the following form

$$\left\{ \begin{array}{lcl} \xi_i(k+1) & = & f_i(\xi_i(k), u_i(k), k) + \Delta f_i(\xi_i(k), k) + w_i(k), \\ y_i(k) & = & h_i(\xi_i(k), \xi_j(k)) + v_i(k) \end{array} \right. \tag{2}$$

where $k \in \{0, 1, 2, \cdots\}$ is the discrete time index. In this project, we deal with the general class of multiagent dynamical systems, and the agent dynamics may assume either the continuous-time model in (1) or the discrete-time model in (2).

To achieve multiagent systems coordination, it is necessary that the agents in the group are capable of exchanging information through the sensing/communication networks. For agent i, its output and measurement vector y_i reflects its interaction with other agents ξ_j in the group through communication/sensor channels. In addition, we define a *coordination variable* $x_i = \chi_i(y_i)$ to generically describe the coordination tasks for multiagent systems, where $\chi_i : \Re^{p_i} \mapsto \Re^q$ is a continuous and differentiable function of y_i. By introducing x_i, various coordination tasks such as consensus, rendezvous, cooperative target localization, mobile agents coverage control, distributed resource allocation, and formation control may be embedded into the definition of function $\chi_i(y_i)$. To this end, the multiagent task coordination to be addressed in this project can be recast as cooperative stability issues as defined below.

Definition 1 *Multiagent systems (1) or (2) are said to be cooperative if* $\lim_{t\to\infty}[x_i(t) - x_j(t)] = \mathbf{1}_q 0$, *where* $\mathbf{1}_q$ *is* $q-$*dimensional column vector with all its elements being 1. Multiagent systems (1) or (2) are said to be cooperatively stable (i.e., cooperative and all the state variables of the systems are uniformly bounded) if, for some steady state* $x^{ss} \in \Re^q$, $\lim_{t\to\infty} x_i(t) = x^{ss}$.

As seen in definition 1, the steady state x^{ss} represents the convergence value of the *coordination variables* $x_i(t)$ for all agents in the group. For example, if the coordination task for mulitagent systems (1) is to seek the average consensus, then $x^{ss} = \sum_{i=1}^{N} x_i(0)/N$.

Now, let us define the *objective function* $U_i(x, u_i, t)$ for agent i to accommodate the optimal performance of the multiagent systems coordination. Let the *objective function* $U_i(x, u_i, t)$ be

$$U_i(x, u_i, t) = \int_t^{\infty} L_i(x(\tau), u_i(\tau))d\tau, \tag{3}$$

where $x = [x_1^T, x_2^T, \cdots, x_N^T]^T$ is the stacked overall *coordination variable*, $L_i(\cdot)$ and $\psi(\cdot)$ are the running cost functions. To this end, the multiagent systems task coordination problem studied in this project can be generically described as follows.

Problem 1 *For a network of mulitagent dynamical systems (1) or (2), design cooperatively stabilizing control protocols* $u_i(t)$ *of the form*

$$u_i(t) = \alpha_i(x_i, x_{j_1}, \cdots, x_{j_l}, t), \tag{4}$$

while solving the following optimization problem

$$\mathit{min} \sum_{i=1}^{n} U_i(x) \tag{5}$$

where $x_{j_k}, j_k \in \mathcal{N}_i$ *are the coordination variables of the neighboring agents of agent* i*,* $\mathcal{N}_i$ *is the index set of the neighboring agents of agent* i*.*

3.2 Sensing/Communication Model and Procedures

The success of solving coordination **Problem 1** is dependent on information exchange among agents. In general, we assume that the information exchange among agents are done through communication broadcasting or agents' sensing capabilities. We consider flexible time-varying sensing/communication topologies among agents. To precisely account for the sensing/communication information exchange among agents in the design of coordination strategy and control protocols, we introduce the following sensing /communication matrix and its corresponding time sequence $\{t_k^s : k = 0, 1, \cdots\}$. That is, within time interval $[t_k^s, t_{k+1}^s)$, the sensing/communication topology is assumed to be unchanged.

$$S(t) = \begin{bmatrix} s_{11} & s_{12}(t) & \cdots & s_{1q}(t) \\ s_{21}(t) & s_{22} & \cdots & s_{2q}(t) \\ \vdots & \vdots & \ddots & \vdots \\ s_{q1}(t) & s_{q2}(t) & \cdots & s_{qq} \end{bmatrix}, \tag{6}$$

with $S(t) = S(t_k^s), \forall t \in [t_k^s, t_{k1}^s)$, where $s_{ii} \equiv 1$; $s_{ij}(t) = 1$ if the jth agent is in the sensor/communication range of the ith agent at time t, and $s_{ij} = 0$ otherwise; and $t_0^s \stackrel{\triangle}{=} t_0$. It can be assumed without loss of generality that $0 < \underline{c}_t \leq t_{k+1}^s - t_k^s \leq \overline{c}_t < \infty$, where $\underline{c}_t$ and $\overline{c}_t$ are constant bounds. In the presence of communication delay τ, the available information at time instant t will depend on $S(t - \tau)$.

In the following sections, we report several multiagent coordination control algorithms in solving **Problem 1** by focusing on procedures of dealing with the following key elements.

- First, a sensing/communication model is fundamental to describe the information exchange among multiple agents in the system. One of the main objectives of this project is to establish the least restrictive network controllability condition for multiagent systems to achieve the task coordination.
- Second, in order to cover a broad class of practical applications for multiple agents, multiagent dynamics are of paramount importance in the coordination tasks. We design optimal and adaptive coordination controls for a general class of dynamical systems with uncertainties.
- Third, multiagent task coordination applications are conducted by particularly solving the formation control and coverage control problems for multiple mobile robots with kinematic constraints.
- Forth, extensive computer simulation and experimental tests have been performed to illustrate the proposed designs.

4.0 RESULTS AND DISCUSSION: NETWORK CONTROLLABILITY ANALYSIS

In this section, we report the results on network controllability analysis and present a new discontinuous cooperative control for consensus of multiagent systems under directed and switching sensing/communication topologies and time-delays. Simulation test results for underwater sonar data transmission are also given.

4.1 Sequential Completeness Condition on Network Controllability

One of the key issues in engineered multiagent systems is the study of network controllability. The objective is to figure out the connectivity conditions on sensor/communication topologies of the network for achieving consensus behavior. In [7][20], the condition is obtained for composite undirected graphs which need to be connected. Extensions were made in [17][9] to the case with directed graphs, and the less restrictive conditions are stated as that there exists a spanning tree or the network is strongly connected periodically. Complement to the aforementioned graph-theoretical

methods, a matrix-theoretical framework is developed in [16] to deal with the high-order systems with arbitrary but finite relative degrees. The notion of *sequentially completeness* was introduced in [16][29] to describe the least required condition on network connectivity for cooperative control design, which is restated by the following definitions.

Definition 2 *Sensing/communication matrix sequence $\{S(t)\}$ is said to be sequentially lower-triangularly complete if it is sequentially lower-triangular and in every row i of its lower triangular canonical form, there is at lease one $j < i$ such that the corresponding block $S_{ij}(t)$ is uniformly non-vanishing.*

Definition 3 *Sensing/communication matrix sequence $\{S(t)\}$ is said to be sequentially complete if the sequence contains an infinite subsequence that is sequentially lower-triangularly complete.*

As an example for sequential completeness, let us assume that the sensing/communication topologies for 3 agents are changing according to the sequence $\{S(t_k), k \in \aleph, \aleph \triangleq \{1, 2, \dots\}\}$ defined below: $S(t_k) = S_1$ for $k = 4\eta$, $S(t_k) = S_2$ for $k = 4\eta + 1$, $S(t_k) = S_3$ for $k = 4\eta + 2$, and $S(t_k) = S_4$ for $k = 4\eta + 3$, where $\eta \in \aleph$,

$$S_1 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad S_2 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$
$$S_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad \text{and} \quad S_4 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{7}$$

The bitwise union of $S_i, i = 1, \cdots, 4$ is

$$\bigcup_i S_i = \left[\begin{array}{cc:c} 1 & 1 & 0 \\ 1 & 1 & 0 \\ \hdashline 1 & 0 & 1 \end{array}\right] \triangleq \left[\begin{array}{c:c} S'_{\Lambda,11} & \emptyset \\ \hdashline S'_{\Lambda,21} & 1 \end{array}\right].$$

It then follows from the structure of $\bigcup_i S_i$ that the corresponding sequence is lower-triangularly complete, and therefore the switching sensing/communication topologies defined by (7) is *sequentially complete*.

4.2 Multiagent Coordination with Discontinuous Dynamics

4.2.1 Issues with Sequential Completeness Condition in the Presence of Discontinuous Dynamics

Let us consider the consensus problem for the simplest multiagent systems described by single-integrator dynamics

$$\dot{x}_l = u_l, \tag{8}$$

where $l \in \Omega \stackrel{\triangle}{=} \{1, \cdots, n\}$, $x_l(t) \in \Re$ is the state, $u_l \in \Re$ is the control input to be designed. The objective is to design $u_l(t)$ to achieve the consensus of the multiagent system (8), that is,

$$\lim_{t \to \infty} x_l(t) = x^*, \quad \forall l, \tag{9}$$

where x^* is some constant denoting the consensus value.

For the cooperative control of multiagent systems (8), if the standard design of $u_i(t)$ is adopted as given below

$$u_l(t) = \sum_{j=1}^{n} \alpha_{lj}(s_{lj}(t_k^s))(x_j(t) - x_l(t)), t \in [t_k^s, t_{k+1}^s), \tag{10}$$

where

$$\alpha_{lj}(t_k^s) = \frac{s_{lj}(t_k^s)}{\sum_{i=1}^{n} s_{li}(t_k^s)}, \tag{11}$$

then it has been proved in [16] that the sequential completeness of sensing/communication matrix sequence $\{S(t)\}$ is the necessary and sufficient condition for consensus of multiagent systems.

However, in practice it is often necessary to consider the following discontinuos cooperative control of the form

$$u_l(t) = \sum_{j=1}^{n} \alpha_{lj}(s_{lj}(t_k^s), x_j(t_k^s))\text{sgn}(x_j(t) - x_l(t)), t \in [t_k^s, t_{k+1}^s), \tag{12}$$

where $\alpha_{lj}(\cdot, \cdot)$ is a nonlinear gain to be designed based on the sensing/communication topology $S(t_k^s)$ as well as the available boundary values $x_j(t_k^s)$ if $s_{lj}(t_k^s) \neq 0$, and sgn$(\cdot)$ function is defined as

$$\text{sgn}(z) = \begin{cases} 1, & z > 0 \\ 0, & z = 0 \\ -1, & z < 0 \end{cases}$$

Remark 1 *By including sgn$(\cdot)$ in the control law (12), it may provide benefits to deal with the control of truly nonlinear systems such as nonholonomic mobile robots [5], and improve the convergence speed of the system. In addition, the control (12) may reduce the sensing/communication loads because on one hand nonlinear gain α_{lj} only relies on states $x_j(t_k^s)$ at the time instants whenever the communication topology changes, and on the other hand, information exchange of sgn$(x_j(t) - x_l(t))$ may also significantly reduce the required transmission capacity compared with that of $(x_j(t) - x_l(t))$.* $\diamond$

Under cooperative control (12), the closed-loop multiagent system become system with discontinuous dynamics. The sequential completeness of sensing/communication network may no longer ensure the consensus if the gains α_{il} are simply designed using (11). This is illustrated through the following example.

Example 1 *Suppose we have 3 agents. Define index set* $\Omega = \{1,2,3\}$, $\Omega_{\max} = \{i \in \Omega : x_i(t) = x_{\max}(t) \stackrel{\triangle}{=} \max_j x_j(t)\}$, *and* $\Omega_{\min} = \{i \in \Omega : x_i(t) = x_{\min}(t) \stackrel{\triangle}{=} \min_j x_j(t)\}$.

Assume that at time instant t_0, *we have* $\Omega_{\min}(t_0) = \{1\}$, *and* $\Omega_{\max}(t_0) = \{2,3\}$, *and the sensing/communication topologies among three agents switch according to sensing/communication matrices* $S(t_{3k}), S(t_{3k+1})$ *and* $S(t_{3k+2})$ *defined below.*

$$S(t_{3k}) = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, S(t_{3k+1}) = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, S(t_{3k+2}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix},$$

where $k = 0, 1, \cdots$. *It can be readily verified that the matrix sequence* $S(t_{3k}), S(t_{3k+1}), S(t_{3k+2})$ *is sequentially complete. However, the consensus is not guaranteed if the standard gain design for* α_{ij} *in (11) is applied under control (12). One possible scenario is that according to the sensing/communication matrix* $S(t_0)$, *agent* 2 *receives information from agent* 1 *and may converge to agent* 1 *in finite time interval* $t_1 - t_0$, *thus at time instant* t_1, *we could have* $\Omega_{\min}(t_1) = \{1,2\}$ *and* $\Omega_{\max}(t_1) = \{3\}$*; similarly, according to* $S(t_1)$, *agent* 1 *receives information from agent* 3 *and may converge to agent* 3 *in finite time* $t_2 - t_1$, *thus we may have* $\Omega_{\min}(t_2) = \{2\}$ *and* $\Omega_{\max}(t_2) = \{1,3\}$*; by* $S(t_2)$, *agent* 3 *receives information from agent* 2 *and may converge to agent* 2 *in finite time* $t_3 - t_2$, *and we may have* $\Omega_{\min}(t_3) = \{2,3\}$ *and* $\Omega_{\max}(t_3) = \{1\}$. *This pattern will repeat following the periodical sensing/communication matrix sequence* $\{S(t_i)\}$. *In other words, though within time interval* $[t_0, t_3)$, *the communication topology is complete, contraction mapping is not established since we have* $x_{\max}(t_3) = x_{\max}(t_0)$ *and* $x_{\min}(t_3) = x_{\min}(t_0)$ *from the above analysis. This is further illustrated in figure 1, in which we consider three agents with controls (12) and gain* $\alpha_{ij}(t)$ *are chosen based on (11), simulation parameters are given as* $t_{3k+i} - t_{3k+i-1} = 0.1, i = 1, 2, k = 0, 1, \cdots$, *and initial conditions* $x_1(t_0) = 0.01, x_2(t_0) = -0.01$, *and* $x_3(t_0) = 0.1$. *Apparently, no consensus is reached.* ◇

4.2.2 Design of New Discontinuous Cooperative Controls

As shown in examples 1, standard network topology based control gain design for (12) no longer implies the consensus of multiagent systems, even with the most-restrictive network connectivity condition (that is, fixed and undirected communication). In this subsection, in order to ensure the multiagent systems consensus with control (12) under the least restrictive sensing/communication condition (that is, sequential completeness of $\{S(t_k^s\}$), we propose a new nonlinear piecewise gain design. The convergence of the overall closed-loop systems is proved by developing a contraction mapping method for multiagent systems.

Theorem 1 *Consider the multiagent system (8) under cooperative control (12). Assume that sensing/communication matrix sequence* $\{S(t_k^s)\}$ *is uniformly sequentially complete*. Let the nonlinear gain* α_{lj} *be designed as follows: for any agent l,*

*The time-varying sensing/communication topology is considered here. If the topology becomes fixed after certain time, we can treat it as a special case of switching sensing/communication sequence $S(t_k^s)$ with $\bar{c}_t$ being any positive constant.

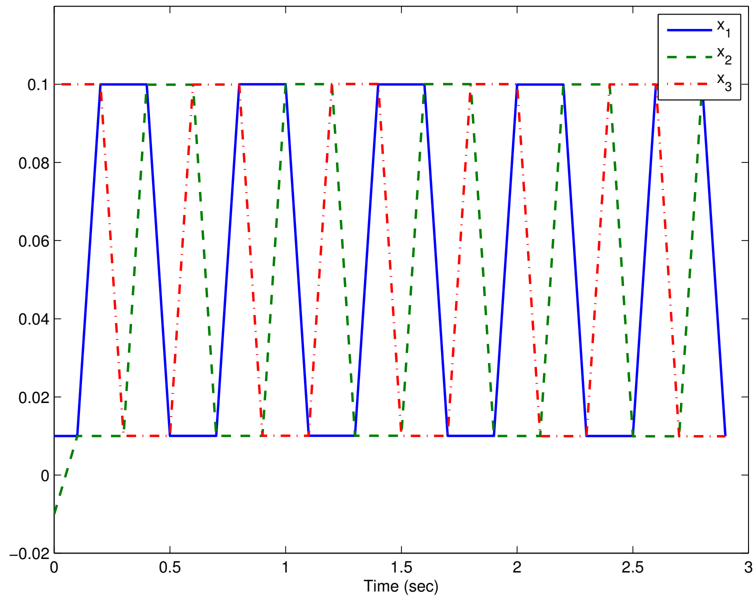

Figure 1: System responses

1) if $x_l(t_k^s) = \max_{j\in\mathcal{N}_l} x_j(t_k^s) = \min_{j\in\mathcal{N}_l} x_j(t_k^s)$*, then* $\alpha_{lj}(t_k^s)$ *can be any bounded positive value.*

2) if $x_l(t_k^s) \geq \max_{j\in\mathcal{N}_l} x_j(t_k^s)$*, let* $\alpha_{lj}(t_k^s)$ *be selected to satisfy the inequality*

$$0 \leq \sum_{j\in\mathcal{N}_l} \alpha_{lj}(t_k^s) < \frac{x_l(t_k^s) - \min_{j\in\mathcal{N}_l} x_j(t_k^s)}{\bar{c}_t} \tag{13}$$

3) if $x_l(t_k^s) \leq \min_{j\in\mathcal{N}_l} x_j(t_k^s)$*, let* $\alpha_{lj}(t_k^s)$ *be selected to satisfy the inequality*

$$0 \leq \sum_{j\in\mathcal{N}_l} \alpha_{lj}(t_k^s) < \frac{\max_{j\in\mathcal{N}_l} x_j(t_k^s) - x_l(t_k^s)}{\bar{c}_t} \tag{14}$$

4) if $\min_{j\in\mathcal{N}_l} x_j(t_k^s) < x_l(t_k^s) < \max_{j\in\mathcal{N}_l} x_j(t_k^s)$*, let* $\alpha_{lj}(t_k^s)$ *be selected to satisfy the inequality*

$$0 \leq \sum_{j\in\mathcal{N}_l} \alpha_{lj}(t_k^s) < \min\left(\begin{array}{c} \frac{\max_{j\in\mathcal{N}_l} x_j(t_k^s) - x_l(t_k^s)}{\bar{c}_t}, \\ \frac{x_l(t_k^s) - \min_{j\in\mathcal{N}_l} x_j(t_k^s)}{\bar{c}_t} \end{array} \right) \tag{15}$$

Then consensus of system (8) is asymptotically achieved.

Proof: See [26].

Theorem 1 provides a general set of sufficient gain design conditions for the asymptotical stability of discontinuous multiagent systems with directed and switching sensing/communication topologies. The nonlinear gain design conditions (13) to (15) are imposed for the purpose of avoiding the possible states oscillation due to the finite time state reachability of dynamical systems driven by discontinuous functions under certain communication topologies.

The result in theorem 1 can be extended to the case with communication delays. That is, In the presence of sensing/communication delays, the cooperative control in (12) becomes

$$u_l(t) = \sum_{j=1}^{n} \alpha_{lj}(s_{lj}, x_j(t_k^s - \tau_{lj}))\text{sgn}(x_j(t - \tau_{lj}) - x_l(t)), t \in [t_k^s, t_{k+1}^s), \tag{16}$$

where $\tau_{lj} \in [0, r]$ are time delays incurred during transmission with r being the upper bound on latencies of information transmission over the network. The following theorem is in the sequel.

Theorem 2 *[26] Consider the multiagent system (8) under cooperative control (16). Assume that sensing/communication matrix sequence of $\{S(t_k^s)\}$ is sequentially complete. Let the nonlinear gain α_{lj} be designed as follows: for any agent l,*

1) *if $x_l(t_k^s) = \max_{j\in\mathcal{N}_l} x_j(t_k^s - \tau_{lj}) = \min_{j\in\mathcal{N}_l} x_j(t_k^s - \tau_{lj})$, then $\alpha_{lj}(t_k^s)$ can be any bounded positive values.*

2) *if $x_l(t_k^s) \geq \max_{j\in\mathcal{N}_l} x_j(t_k^s - \tau_{lj})$, let $\alpha_{lj}(t_k^s)$ be selected to satisfy the inequality*

$$0 \leq \sum_{j\in\mathcal{N}_l} \alpha_{lj}(t_k^s) < \frac{x_l(t_k^s) - \min_{j\in\mathcal{N}_l} x_j(t_k^s - \tau_{lj})}{\max\{\bar{c}_t, r\}} \tag{17}$$

3) *if $x_l(t_k^s) \leq \min_{j\in\mathcal{N}_l} x_j(t_k^s - \tau_{lj})$, let $\alpha_{lj}(t_k^s)$ be selected to satisfy the inequality*

$$0 \leq \sum_{j\in\mathcal{N}_l} \alpha_{lj}(t_k^s) < \frac{\max_{j\in\mathcal{N}_l} x_j(t_k^s - \tau_{lj}) - x_l(t_k^s)}{\max\{\bar{c}_t, r\}} \tag{18}$$

4) *if $\min_{j\in\mathcal{N}_l} x_j(t_k^s - \tau_{lj}) < x_l(t_k^s) < \max_{j\in\mathcal{N}_l} x_j(t_k^s - \tau_{lj})$, let $\alpha_{lj}(t_k^s)$ be selected to satisfy the inequality*

$$0 \leq \sum_{j\in\mathcal{N}_l} \alpha_{lj}(t_k^s) < \min\left(\begin{array}{c} \frac{\max_{j\in\mathcal{N}_l} x_j(t_k^s - \tau_{lj}) - x_l(t_k^s)}{\max\{\bar{c}_t, r\}}, \\ \frac{x_l(t_k^s) - \min_{j\in\mathcal{N}_l} x_j(t_k^s - \tau_{lj})}{\max\{\bar{c}_t, r\}} \end{array}\right) \tag{19}$$

Then consensus of system (8) is asymptotically achieved.

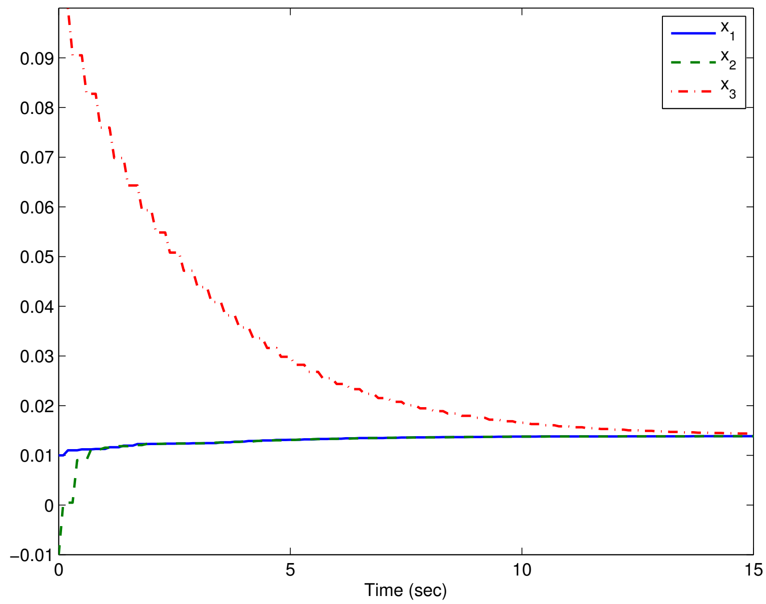

Figure 2: System responses

4.2.3 Simulation Results and Experimental Testing

Let us reconsider example 1 for the consensus of three agents with control (12) under the sensing/communication topologies $S(t_{3k}), S(t_{3k+1})$ and $S(t_{3k+2})$. We choose the nonlinear piecewise constant gain α_{ij} based on theorem 1. Under the same simulation conditions, system responses are shown in figure 2, and consensus is reached.

To further illustrate the benefit of introducing discontinuous dynamics in multiagent systems, the underwater sonar data transmission testing was conducted at the Wave Laboratory, Embry-Riddle Aeronautical University. As shown in figure 3, a wave tank is used as the testbed and wave maker generates the noise environment. Two sonar communication transducers are used for sending and receiving data. 100 sets of data are used in the testing under two scenarios of direct position data transmission and indirect position data transmission. That is, in the scenario of indirect position data transmission, the sign of data (binary number) was transmitted. Figures 4 and 5 show the testing resulting for two scenarios, respectively. The transmitting and receiving speed is $S_5/R_5 = 13bits/s$. It is apparent that the error rate with indirect data transmission is much lower.

5.0 RESULTS AND DISCUSSION: OPTIMAL AND ADAPTIVE MULTI-AGENT COORDINATION

In this section, we present three results on optimal and adaptive multiagent coordination. The first result is on the design of a practically implementable approximately adaptive neural cooperative

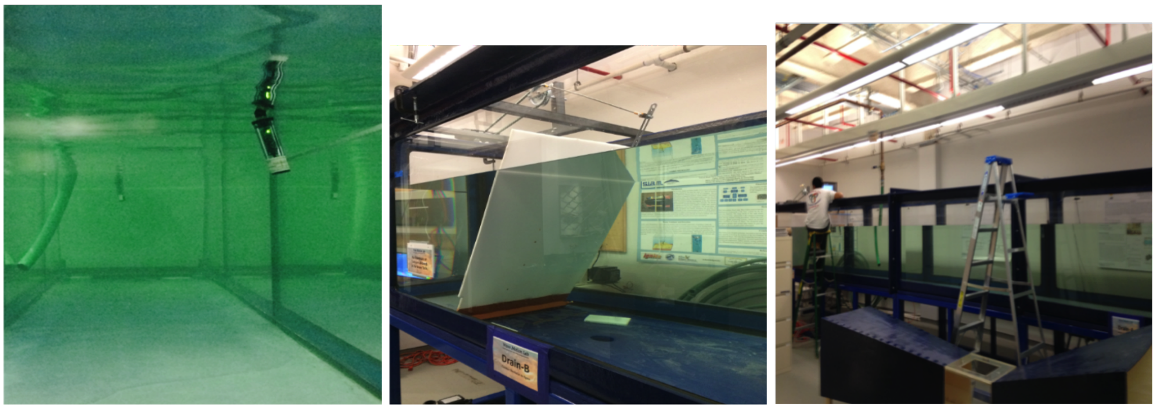

Figure 3: Wave Tank

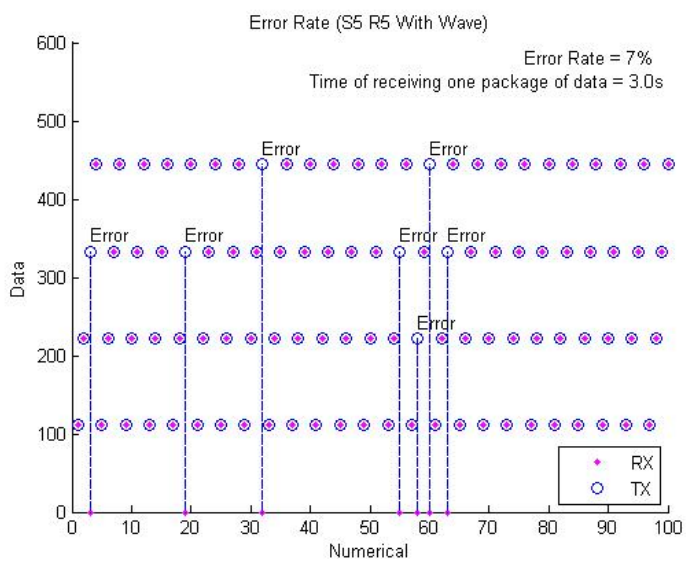

Figure 4: Direct Position Data Transmission

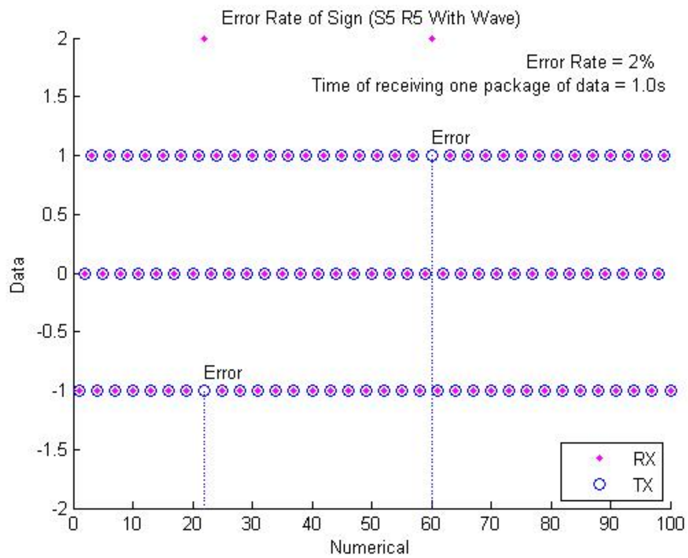

Figure 5: Indirect Position Data Transmission

control for multiagent systems based on online approximate dynamic programming. The second one is on the design of optimal cooperative control based approximate Q-functions. The third one is on a new distributed adaptive cooperative control method for consensus tracking of multiagent systems with model uncertainties.

5.1 Value Function Based Multiagent Policy Iteration

In the study of cooperative control of multiagent systems, fruitful results for cooperative control design have been obtained for first-order linear systems [7, 17, 9, 6, 20], for second-order linear systems [23], for high-order linear systems [16, 28], and for nonlinear systems [11, 10, 12, 14, 4, 29, 26], few results are available for optimal cooperative control design. There appeared some recent work in the study of optimal cooperative control, such as those in [22, 1, 3, 15]. Nonetheless, it is still a challenge to systematically address the optimal cooperative control problem for more general nonlinear multiagent systems, particularly, in the presence of model uncertainties.

The result reported in this section aims to deal with such a challenge. For multiagent optimal cooperative control, the key issue is how to establish an optimality equation and find its solution in real time. We tackle this problem by considering a general class of feedback linearizable nonlinear multiagent systems. We assume that there exist admissible cooperative controls for such kind of multiagent systems under the *complete* sensing/communication condition [16]. The fixed sensing/communication topology is imposed for ease of design. The case for more complicated time-varying sensing communication topology will be treated in future work. The optimal cooperative

control problem is then formulated as making all systems achieve consensus while minimizing the individual sensing/communication topology dependent cost functions. It is shown that the optimal solution to the defined problem requires to solve a multiagent Hamilton-Jacobi-Bellman (HJB) equation. To avoid the obstacles in analytically solving multiagent HJB equation, we extend the online policy iteration approach in [21][25] to the multiagent case, and employ radial basis function (RBF) neural networks to approximate value functions at each iteration. Through seeking the least-squares solution to estimate the optimal neural weights, a new approximately adaptive multiagent policy iteration algorithm is proposed. It is further shown that the proposed adaptive optimal cooperative control approximately solves the posed optimal consensus problem. Simulation results are provided to illustrate the effectiveness of the proposed optimal design.

5.1.1 Problem Formulation

Consider a multiagent system which has N members and each agent assumes the general nonlinear dynamics

$$\dot{x}_i = f_i(x_i) + g_i(x_i)u_i, \tag{20}$$

where $i \in \Omega \stackrel{\triangle}{=} \{1, \cdots, N\}$, $x_i(t) \in \Re^n$ is the system state, $u_i \in \Re^m$ is the control input to be designed, $f_i, g_i : \Re^n \mapsto \Re^n$ are locally Lipschitz continuous functions.

The objective is to design an optimal cooperative control $u_i(t)$ to achieve the consensus of the multiagent system (20) such that

$$\lim_{t\to\infty} x_i(t) = x^*, \quad \forall i, \tag{21}$$

while minimizing the following individual cost function for each agent i,

$$J_i(u_i; x_i(t_0), s_{ij}x_j(t_0)) = \int_{t_0}^{\infty} \left(\sum_{j=1}^{N} (x_i - x_j)^T s_{ij} Q_{ij} (x_i - x_j) + u_i^T R_i u_i \right) dt, \tag{22}$$

where x^* is some constant denoting the consensus value, Q_{ij} and R_i are symmetric and positive definite matrices, and s_{ij} (defined in equation (6)) is a binary number describing the availability of the sensing/communication information exchange between the agent i and the agent j.

It should be noted that the individual cost function defined in (22) is related to the measurement of closeness of x_i to x_j and control effort of agent i. In contrast, a multiagent differential graphical game problem was defined in [24], in which a cooperative performance index was imposed by including the terms related to control efforts of neighboring agents. Here we are not looking into the team performance index as formulated in cooperative game based solution. Instead, we deal with the optimal solution by looking into minimizing individual performance defined in (22).

5.1.2 On the design of Approximately Adaptive Cooperative Optimal Control

Multiagent HJB Equation: The design starts with the development of multiagent HJB equations.

Recall that the cost function for agent i is defined in (22), which may be rewritten as

$$J_i(u_i; x_i(t_0), s_{ij}x_j(t_0)) = \int_{t_0}^{\infty} \left(\sum_{j \in \mathcal{N}_i} (x_i - x_j)^T Q_{ij}(x_i - x_j) + u_i^T R_i u_i \right) dt, \tag{23}$$

where $\mathcal{N}_i = \{j \in \Omega | s_{ij} \neq 0\}$ denotes the neighbor set of agent i. The following lemma is instrumental in developing the multiagent Hamilton-Jacobi-Bellman (HJB) equation.

Lemma 1 *For admissible cooperative control $u_i(t)$, if there exists a positive definite continuously differentiable function $V_i(x_i, s_{ij}x_j; u_i)$ satisfying the following property*

$$\begin{aligned} &\frac{\partial V_i^T}{\partial x_i}(f(x_i) + g(x_i)u_i) + \sum_{j \in \mathcal{N}_j} \frac{\partial V_j^T}{\partial x_j}(f(x_j) + g(x_j)u_j) \\ &+ \sum_{j \in \mathcal{N}_i} (x_j - x_i)^T Q_{ij}(x_j - x_i) + u_i^T R_i u_i = 0 \end{aligned} \tag{24}$$

and the boundary condition $V_i(x_i(\infty), s_{ij}x_j(\infty); u_i) = 0$, then $V_i(x_i, s_{ij}x_j; u_i)$ is the value function for system (20) for all t, and

$$V_i(x_i(t_0), s_{ij}x_j(t_0); u_i) = J_i(u_i; x_i(t_0), s_{ij}x_j(t_0)) \tag{25}$$

It follows from lemma 1 and Bellman's principle of optimality, we know that the optimal value function $V_i^*(x_i(t), s_{ij}x_j(t))$ approximately satisfies for small $\Delta \to 0$

$$V_i^*(x_i(t), s_{ij}x_j(t)) \simeq \min_{u_i}[l(x_i(t), x_j(t), u_i)\Delta + V_i^*(x_i(t+\Delta), s_{ij}x_j(t+\Delta)], \tag{26}$$

where $l(x_i(t), x_j(t), u_i) \triangleq \sum_{j \in \mathcal{N}_i}(x_i - x_j)^T Q_{ij}(x_i - x_j) + u_i^T R_i u_i$, $x_i(t+\Delta) \simeq x_i(t) + (f(x_i) + g(x_i)u_i)\Delta$, and $x_j(t+\Delta) \simeq x_j(t) + (f(x_j) + g(x_j)u_j)\Delta$.

The corresponding optimal cooperative control can be derived as

$$u_i^* = -\frac{1}{2} R_i^{-1} g_i^T \frac{\partial V_i^*}{\partial x_i} \tag{27}$$

and the *mutliagent HJB equation* is

$$\begin{aligned} 0 \;=\; &\sum_{j \in \mathcal{N}_i} (x_j - x_i)^T Q_{ij}(x_j - x_i) + \frac{\partial V_i^{*T}}{\partial x_i} f(x_i) - \frac{1}{4} \frac{\partial V_i^{*T}}{\partial x_i} g(x_i) R_i^{-1} g(x_i)^T \frac{\partial V_i^*}{\partial x_i} \\ &+ \sum_{j \in \mathcal{N}_i} \frac{\partial V_i^{*T}}{\partial x_i}(f(x_j) + g(x_j)u_j), \end{aligned} \tag{28}$$

with the associated boundary condition $V_i^*(x_i^*, s_{ij}x_j^*) = 0$, which requires that the optimal value must be null when evaluated on an extremal trajectory (all agents in the set $\{i, \mathcal{N}_i\}$ reach consensus.)

The solution to (28) would provide the optimal cooperative control in (27). However, it is difficult to solve mainly for two reasons. First, equation (28) is a nonlinear partial differential equation, and it is in general impossible to solve this equation in analytic form. Second, the coupling terms $\sum_{j\in\mathcal{N}_i} \frac{\partial V_i^{*T}}{\partial x_i}(f(x_j) + g(x_j)u_j)$ cause extra difficulty due to involvement of u_j which may require information propagation from agents not in the neighboring set $\mathcal{N}_i$.

The Proposed Multiagent Policy Iteration Algorithm: The proposed multiagent policy iteration algorithm consists of the following two steps:

Step 1: **Policy evaluation**. Find an admissible cooperative control policy $u_{i,0}(x_i, s_{ij}x_j)$. For any integer $l \geq 0$ denoting the iteration index, solve for $V_{i,l}(x_i, s_{ij}x_j; u_{i,l})$ using

$$\begin{aligned} 0 \;=\; & \sum_{j\in\mathcal{N}_i}(x_j - x_i)^T Q_{ij}(x_j - x_i) + u_{i,l}^T R_i u_{i,l} + \frac{\partial V_{i,l}^T}{\partial x_i}(f(x_i) + g(x_i)u_{i,l}) \\ & + \sum_{j\in\mathcal{N}_i}\frac{\partial V_{i,l}^T}{\partial x_j}(f(x_j) + g(x_j)u_{j,l}), \end{aligned} \tag{29}$$

with $V_{i,l}(x^*, s_{ij}x^*) = 0$.

Step 2: **Policy improvement**. Update the control policy by

$$u_{i,l+1} = -\frac{1}{2}R_i^{-1}g_i^T\frac{\partial V_{i,l}}{\partial x_i} \tag{30}$$

The convergence of the multiagent policy iteration algorithm given in (29) and (30) is summarized into the following theorem.

Theorem 3 *If a sequence of pairs $\{V_{i,l}, u_{i,l+1}\}$ is generated by (29) and (30), then the corresponding value function $V_{i,l}$ satisfying*

$$V_{i,l+1} \leq V_{i,l} \tag{31}$$

and

$$\lim_{l\to\infty} V_{i,l} = V_i^* \tag{32}$$

Proof: See [31].

For the ease of implementation, the **policy evaluation** step in the proposed multiagent policy iteration algorithm can be replaced by following equation (33) for solving for $V_{i,l}$ based on the

available information $x_i(t), x_j(t)$ and $u_{i,l}(t)$ during the time interval $[t, t+T]$.

$$V_{i,l}(x_i(t), s_{ij}x_j(t); u_{i,l}(t)) = \int_t^{t+T} \left(\sum_{j\in\mathcal{N}_i} (x_j - x_i)^T Q_{ij}(x_j - x_i) + u_{i,l}^T R_i u_{i,l} \right) dt + V_{i,l}(x_i(t+T), s_{ij}x_j(t+T); u_{i,l}(t+T)), \quad (33)$$

where $T > 0$ represents certain time interval.

Approximately Adaptive Cooperative Optimal Control: A significant advantage of the proposed multiagent policy iteration algorithm is that it iteratively generates a sequence of pairs $\{V_{i,l}, u_{i,l+1}\}$ through (33) and (30) by only using the available local information x_i, x_j and u_i for agent i, which monotonically converges to the optimal value V_i^* and u_i^*. It is apparent that the key is to solve for $V_{i,l}$ from (33). For the unknown value functions $V_{i,l}(x_i, s_{ij}x_j)$, we use the following neural network approximator.

$$V_{i,l}(x_i, s_{ij}x_j) = \Phi_{i,l}^T(\bar{x}_i)\theta_{i,l}^* + \omega_{i,l}(\bar{x}_i), \ \ \forall \bar{x}_i \in \Omega_i \quad (34)$$

where $\bar{x}_i = [s_{i1}x_1, s_{i2}x_2, \cdots, x_i, \cdots, s_{ij}x_j, \cdots, s_{iN}x_N]^T$, $\theta_{i,l}^* \in R^{l_i}$ is an unknown constant parameter vector, the neural network node number $l_i > 1$, $\omega_{i,l}(\bar{x}_i)$ is the approximation error, and $\Phi_{i,l}(\bar{x}_i) = [\phi_{i_1,l}, \phi_{i_2,l}, \cdots, \phi_{i_{l_i},l}]^T$ is the known basis function vector.

Upon using the function approximator (34), the policy evaluation equation in (33) becomes

$$\int_t^{t+T} \left(\sum_{j\in\mathcal{N}_i} (x_j - x_i)^T Q_{ij}(x_j - x_i) + u_{i,l}^T R_i u_{i,l} \right) dt = [\Phi_{i,l}(\bar{x}_i(t)) - \Phi_{i,l}(\bar{x}_i(t+T))]^T \theta_{i,l}^* + \bar{\omega}_{i,l}(t), \quad (35)$$

where $\bar{\omega}_{i,l}(t) = \omega_{i,l}(t) - \omega_{i,l}(t+T)$.

It follows from (35) that $\theta_{i,l}^*$ provides the best approximate solution for the policy evaluation. However, its value is unknown, and needs to identified online. Let $\theta_{i,l}(t)$ be the estimate of $\theta_{i,l}^*$, and equation (35) becomes

$$\int_t^{t+T} \left(\sum_{j\in\mathcal{N}_i} (x_j - x_i)^T Q_{ij}(x_j - x_i) + u_{i,l}^T R_i u_{i,l} \right) dt = [\Phi_{i,l}(\bar{x}_i(t)) - \Phi_{i,l}(\bar{x}_i(t+T))]^T \theta_{i,l}(t) + e_{i,l}(t), \quad (36)$$

where $e_{i,l}(t) = [\Phi_{i,l}(\bar{x}_i(t)) - \Phi_{i,l}(\bar{x}_i(t+T))]^T \tilde{\theta}_{i,l}(t) + \bar{\omega}_{i,l}(t)$, $\tilde{\theta}_{i,l}(t) = \theta_{i,l}^* - \theta_{i,l}(t)$. Thus, given any admissible cooperative control, the parameter $\theta_{i,l}$ should be chosen to minimize the squared approximation residual error $e_{i,l}^2(t)$. As $\theta_{i,l}(t) \to \theta_{i,l}^*$, it is obvious that $e_{i,l}(t) \to \bar{\omega}_{i,l}$.

In what follows, we present the proposed adaptive law for $\theta_{i,l}$ using the least-squares estimation. Let us define

$$z_i(t_k) = \int_{t_k}^{t_{k+1}} \left(\sum_{j\in\mathcal{N}_i} (x_j - x_i)^T Q_{ij}(x_j - x_i) + u_{i,l}^T R_i u_{i,l} \right) dt$$

and

$$\Psi_{i,l}(t_k) = \Phi_{i,l}(\bar{x}_i(t_k)) - \Phi_{i,l}(\bar{x}_i(t_{k+1})).$$

Substituting this into (36) yields

$$z_i(t_k) = \Psi_{i,l}(t_k)^T \theta_{i,l} + e_{i,l}(t_k) \tag{37}$$

The model in (37) is the *regression model* for policy iteration and $\Psi_{i,l}$ is called the *regressor*. Through the observation interval $[t_k, t_{k+n}]$, pairs of observations and regressors $\{(z_i(t_{k+\mu}), \Psi_{i,l}(t_{k+\mu})), \mu = 0, 1, n-1\}$ are obtained by using control policy $u_{i,l}$. The parameter $\theta_{i,l}$ will be chosen to minimize the least-squares loss function

$$L(\theta_{i,l}, t_k) = \frac{1}{2} \sum_{\mu=1}^{n-1} \left(z_i(t_{k+\mu}) - \Psi_{i,l}(t_{k+\mu})^T \theta_{i,l} \right)^2 .$$

To this end, standard least-squares estimation algorithm renders

$$\theta_{i,l} = \left(\Xi_{i,l}^T \Xi_{i,l} \right)^{-1} \Xi_{i,l}^T Z_{i,l} \tag{38}$$

where $Z_{i,l} = [z_i(t_k), z_i(t_{k+1}), \cdots, z_i(t_{k+n-1})]^T$, and $\Xi_{i,l} = [\Psi_{i,l}^T(t_k), \cdots, \Psi_{i,l}^T(t_{k+n-1})]^T$. Thus, according to **policy improvement** step in (30), and noting $\frac{\partial V_{i,l}}{\partial x_i} = \frac{\partial \Phi_{i,l}^T}{\partial x_i} \theta_{i,l}$ the control law is

$$u_{i,l+1} = -\frac{1}{2} R_i^{-1} g_i^T \frac{\partial \Phi_{i,l}^T}{\partial x_i} \theta_{i,l}. \tag{39}$$

The above results can be summarized into the following proposition.

Proposition 1 *Under assumption of complete sensing/communication topology, the control law (39) with adaptive law (38) approximately solves the optimal cooperative consensus problem for multiagent nonlinear system (20) by minimizing the cost function (22).*

In summary, the proposed approximately adaptive multiagent policy iteration (MPI) algorithm is given in Algorithm 1.

Algorithm 1 Approximately Adaptive MPI Algorithm

1: Let $l = 0$. Given initial states $x_i(t_0)$, $s_{ij} x_j(t_0)$, let the initial admissible cooperative control policy be $u_{i,0}$.

2: Employ the control policy $u_{i,l}$, and within the observation interval $[t_{l\times n}, t_{(l+1)n-1}]$, collect the data pairs

$$\{(z_i(t_{l\times n+\mu}), \Psi_{i,0}(t_{l\times n+\mu})), \mu = 0, 1, n-1\}$$

3: Solve for $\theta_{i,l}$ from (38).

4: Solve for $u_{i,l+1}$ from (39).

5: Let $l \leftarrow l + 1$, and go to step 2, until

$$\|\theta_{i,l+1} - \theta_{i,l}\|^2 \leq \epsilon,$$

where $\epsilon > 0$ is a sufficiently small predefined threshold.

5.1.3 Simulation Results

To illustrate the proposed approximately adaptive cooperative control, we consider a simple multiagent system with 3 agents modeled by the following single integrators

$$\dot{x}_i = u_i, \quad i = 1, 2, 3 \tag{40}$$

where $x_i \in \Re$, and $u_i \in \Re$. Let the sensing/communication topology among 3 agents be given by

$$S = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

Apparently, S matrix is complete, and admissible cooperative control exists for the consensus of (40). Select the weight matrices in (22) as $Q_{ij} = 1, R_i = 0.25$ for simulation purpose. We use a single neural node approximator for each value function $V_{i,l}$. Based on S matrix, we choose the basic functions as $\Phi_{1,l} = (x_1 - x_2)^2$, $\Phi_{2,l} = (x_2 - x_3)^2$ and $\Phi_{3,l} = (x_3 - x_1)^2$ for value functions $V_{1,l}, V_{2,l}$ and $V_{3,l}$, respectively. System initial states are $x_1(0) = 0.5, x_2(0) = 0.2$ and $x_3(0) = 0.3$. Figure 6 shows that system states consensus is achieved, figure 7 displays the instantaneous cost values. Figure 8 illustrates the optimal cooperative control inputs, and the convergence of neural network weights estimates is shown in figure 9.

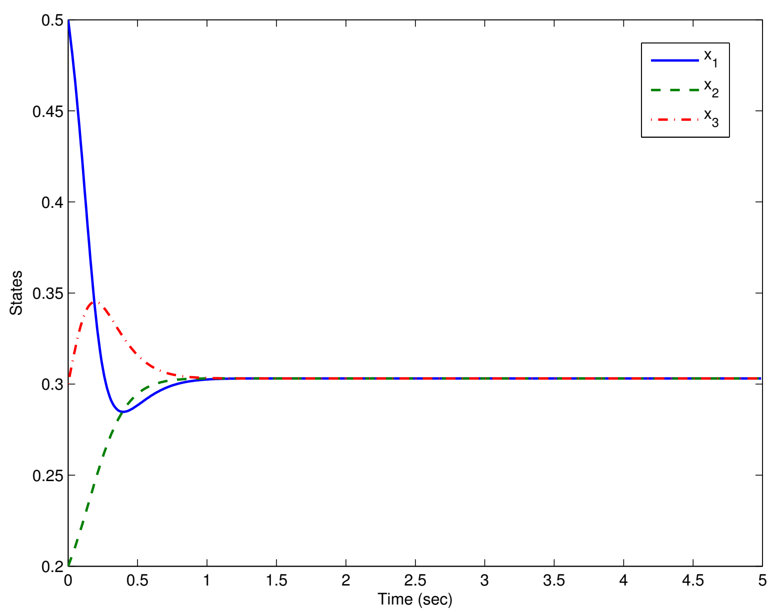

Figure 6: Consensus of $x_i(t)$

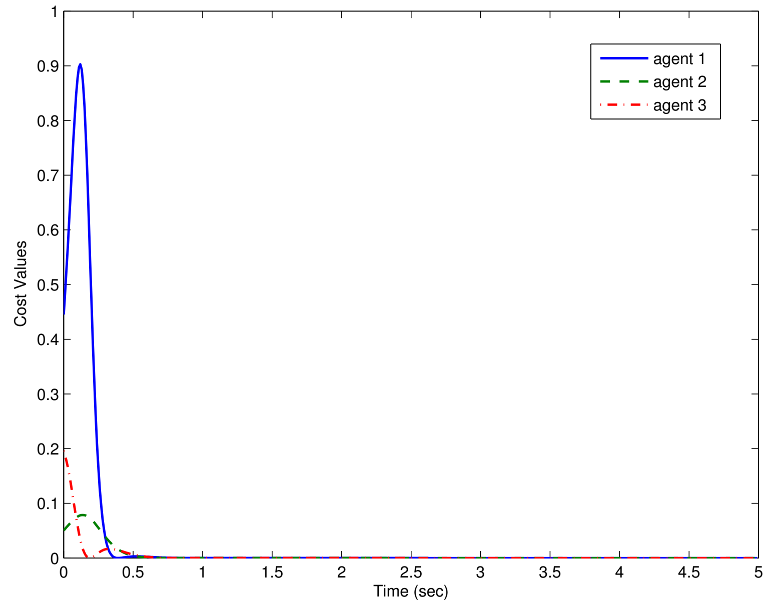

Figure 7: Instantaneous cost values versus time

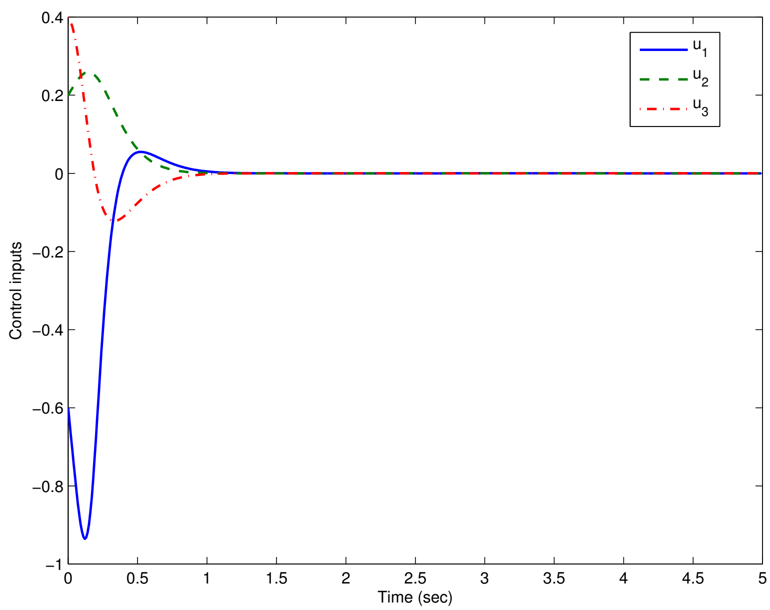

Figure 8: Optimal cooperative controls

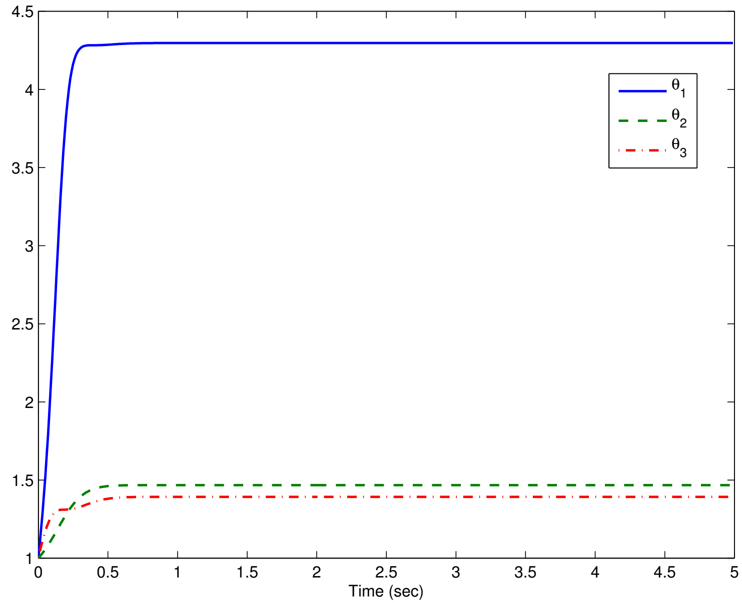

Figure 9: Parameters of neural networks versus time

Simulation was also done for the following three agents with nonlinear models, and the corresponding results in figures 10 to 13 illustrate the performance of the proposed optimal coordination algorithm.

$$
\begin{aligned}
&\text{Agent 1:} \quad \dot{x}_1 = -x_1 + \sin(x_1) + u_1 \\
&\text{Agent 2:} \quad \dot{x}_2 = -x_2 - \sin(x_2) + u_2 \\
&\text{Agent 3:} \quad \dot{x}_3 = u_3
\end{aligned}
$$

5.2 Approximate Q-Function for Multiagent Coordination

There are two main issues with the proposed value function based multiagent policy iteration algorithm. First, the computation of $u_{i,l+1}$ requires to know system dynamics g_i. Second, the implementation of estimation algorithm requires an excitation condition for matrix $\Xi_{i,l}^T \Xi_{i,l}$, which might cause difficult in selecting basis functions for linear approximators. To address those issues, we propose the following new cooperative Q-iteration algorithm.

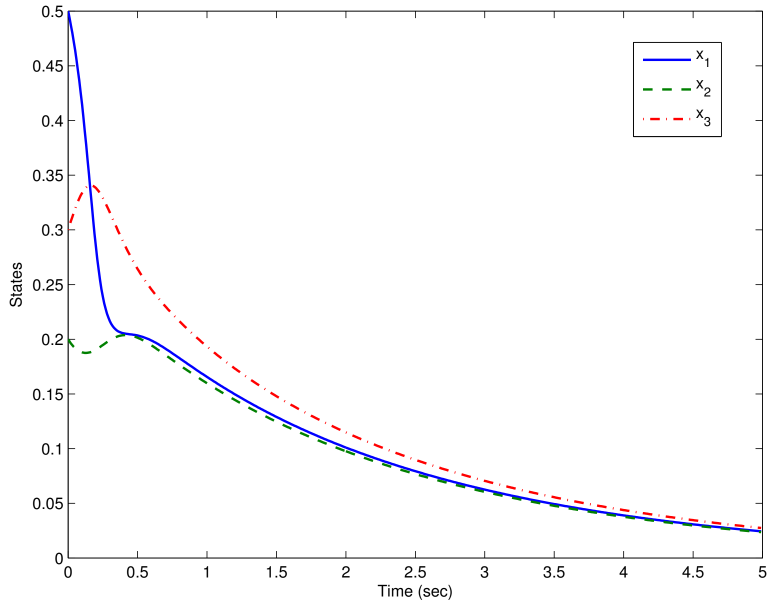

Figure 10: System responses

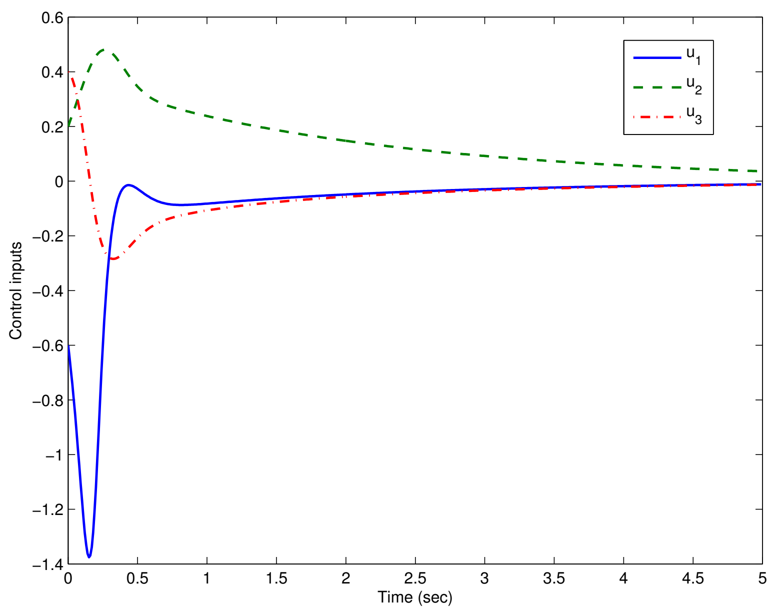

Figure 11: Cooperative control inputs

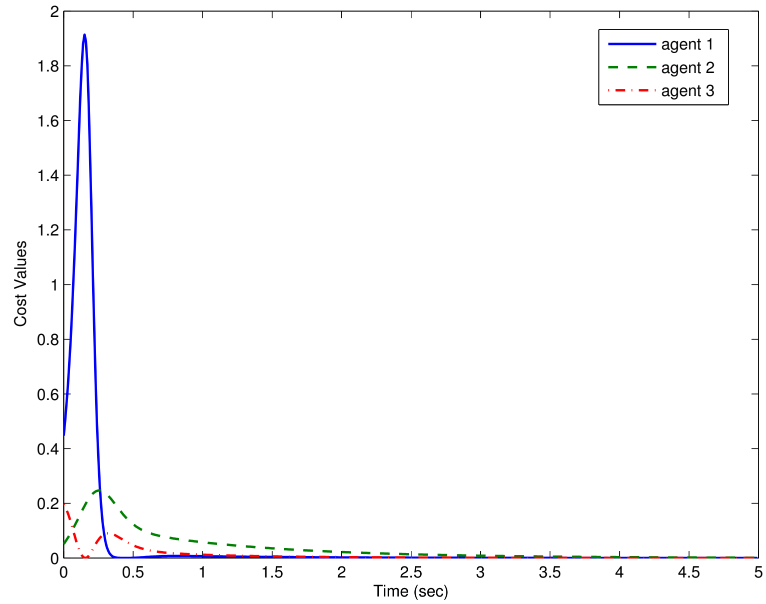

Figure 12: Performance Index

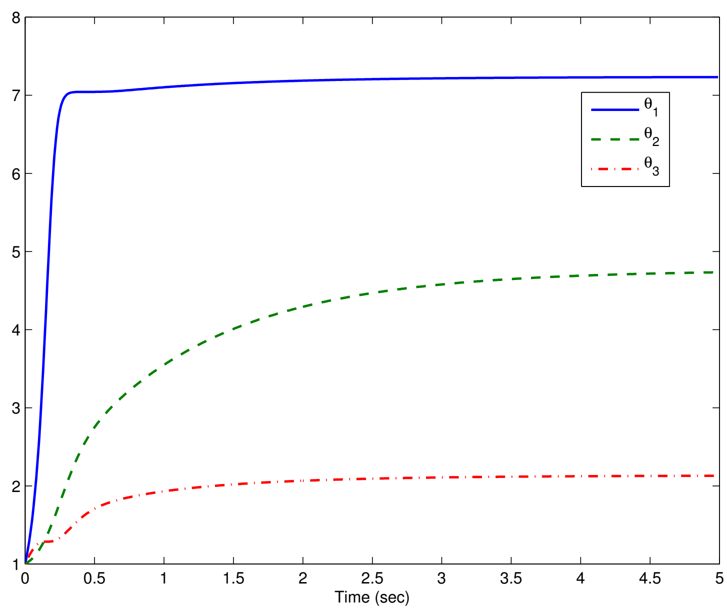

Figure 13: Parameters esimation

5.2.1 Problem Statement

Consider multiagent systems with more general **nonaffine** discrete-time model

$$x_i(k+1) = f_i(x_i(k), u_i(k))$$

The individual cost function for agent i is

$$V_i(x_i(k), s_{ij}x_j(k)) = \sum_{l=k}^{\infty}\left(\sum_{j=1}^{N}(x_i(l) - x_j(l))^T s_{ij}Q_{ij}(x_i(l) - x_j(l)) + u_i^T(l)R_i u_i(l)\right)$$

Q-function is a state-action value function, which gives the value obtained when starting from a given state, applying a given action, and following a policy thereafter. The Q function for agent i is given as follows:

$$\begin{aligned} Q_i(x_i(k), s_{ij}x_j(k), u_i(k)) &= \sum_{j=1}^{N}(x_i(k) - x_j(k))^T s_{ij}Q_{ij}(x_i(k) - x_j(k)) + u_i^T(k)R_i u_i(k) \\ &\quad + V_i(x_i(k+1), s_{ij}x_j(k+1), u_i(k+1)) \end{aligned}$$

Correspondingly, the optimal Q function id defined as

$$\begin{aligned} Q_i^*(x_i(k), s_{ij}x_j(k), u_i(k)) &= \sum_{j=1}^{N}(x_i(k) - x_j(k))^T s_{ij}Q_{ij}(x_i(k) - x_j(k)) + u_i^T(k)R_i u_i(k) \\ &\quad + V_i^*(x_i(k+1), s_{ij}x_j(k+1), u_i(k+1)) \end{aligned}$$

and the optimal cooperative coordination control protocol is

$$u_i^*(x_i(k), s_{ij}x_j(k)) = \text{argmin}_{u_i} Q_i^*(x_i(k), s_{ij}x_j(k), u_i(k))$$

5.2.2 Proposed Coordination Algorithm

To this end, the proposed Q function based approximate multiagent policy iteration algorithm is summarized as two steps.

- **Policy evaluation**
  - According to Bellman equation

$$\begin{aligned} Q_{i,l+1}^+(x_i(k), s_{ij}x_j(k), u_{i,l}(k)) &= \sum_{j=1}^{N}(x_i(k) - x_j(k))^T s_{ij}Q_{ij}(x_i(k) - x_j(k)) \\ &\quad + u_{i,l}^T(k)R_i u_{i,l}(k) + Q_{i,l}(f_i, s_{ij}f_j, u_{i,l}(k)) \end{aligned}$$

– The parameter vector $\theta_{i,l+1}$ is obtained by a projection mapping P

$$\theta_{i,l+1} = P(Q^{+}_{i,l+1})$$

- **Policy improvement**

$$u_{i,l+1} = \text{argmin}_{u_i} Q_{i,l+1}$$

Specifically, by using available on-line data $x_i(k), x_j(k), u_i(k), x_i(k+1), r_i(k+1), u_i(k+1)$, the following temporal difference Q-Iteration can be used in the multiagent Q-function policy iteration algorithm

$$\begin{aligned} Q_{i,k+1}(x_i(k), u_i(k)) &= Q_{i,k}(x_i(k), u_i(k)) + \alpha_k[r_i(k+1) + \gamma Q_{i,k}(x_i(k+1), u_i(k+1)) \\ &\quad - Q_{i,k}(x_i(k), u_i(k))] \end{aligned} \tag{41}$$

However, for systems with large and continuous spaces, Q-functions in terms of state-action pairs will have infinite number of pairs, there is no way to learn and explore all those. Therefore we propose to use the parametric approximation of Q-functions, that is, let

$$Q_i(x_i, u_i) = \phi_i^T(x_i, s_{ij}x_j, u_i)\theta_i$$

where ϕ_i basis functions, θ_i parameter vector to be estimated. To this end, the model-free approximate multiagent Q-learning algorithms can be summarized as follows.

1: Measure initial states $x_i(t_0)$, $s_{ij}x_j(t_0)$.
2: Initialize parameter vector, $\theta_i(0) = 0$.
3: Let $u_i(0) = 0$
4: For every time set $k = 0, 1, 2, \cdots$, do
5: Apply $u_i(k)$, measure $x_i(k+1), x_j(k+1), r_i(k+1)$
6:

$$u_i(k+1) = \frac{\partial \phi_i}{\partial u_i}\theta_i(k)$$

7:

$$\begin{aligned} \theta_i(k+1) &= \theta_i(k) + \alpha_k \left[r_i(k+1) + \gamma \phi_i^T(x_i(k+1), u_i(k+1))\theta_i(k) \right. \\ &\quad \left. - \phi_i^T(x_i(k), u_i(k))\theta_i(k)\right] \phi_i(x_i(k), u_i(k)) \end{aligned}$$

8: end for

5.2.3 Simulation Results

To illustrate the proposed design, we apply the approximate Q-learning algorithm to the consensus control of the following three agents

$$x_i(k+1) = x_i(k) + Tu_i(k), T = 0.05$$

with the Sensing/Communication Topology

$$S = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

The reward function $r_i(k+1)$ are

$$r_1(k+1) = -5(x_2(k) - x_1(k))^2 - 0.01u_1(k)^2$$

$$r_2(k+1) = -5(x_3(k) - x_2(k))^2 - 0.01u_2(k)^2$$

$$r_3(k+1) = -5(x_1(k) - x_3(k))^2 - 0.01u_3(k)^2$$

and the approximate Q-functions

$$Q_1(x_1, u_1) = -((x_2 - x_1)^2 + (x_2 - x_1)u_1 + u_1^2)\theta_1$$

$$Q_1(x_2, u_2) = -((x_3 - x_2)^2 + (x_3 - x_2)u_2 + u_2^2)\theta_2$$

$$Q_1(x_3, u_3) = -((x_1 - x_3)^2 + (x_1 - x_3)u_3 + u_3^2)\theta_3$$

The corresponding cooperative controls are

$$u_1(k) = -\frac{1}{2R_D + 2\theta_1(k)}(x_1(k) - x_2(k))\theta_1(k)$$

$$u_2(k) = -\frac{1}{2R_D + 2\theta_2(k)}(x_2(k) - x_3(k))\theta_2(k)$$

$$u_3(k) = -\frac{1}{2R_D + 2\theta_3(k)}(x_3(k) - x_1(k))\theta_3(k)$$

with adaptation laws

$$\begin{aligned} \theta_i(k+1) \;&=\; \theta_i(k) + \alpha_k \left[r_i(k+1) + \gamma\phi_i^T(x_i(k+1), u_i(k+1))\theta_i(k) \right. \\ &\quad \left. -\phi_i^T(x_i(k), u_i(k))\theta_i(k) \right] \phi_i(x_i(k), u_i(k)) \end{aligned}$$

Simulation parameters: $T = 0.05$; $Q_D = 10$; $R_D = 5$; $\gamma = 0.98$; $\alpha = 0.5$. The simulation results are shown in figures 14 - 17.

5.3 Adaptive Consensus Tracking for Uncertain Multiagent Systems

5.3.1 Problem Statement

Consider the multiagent systems in which the ith agent is described by the scalar differential equation

$$\dot{x}_i = a_i x_i + g_i u_i \tag{42}$$

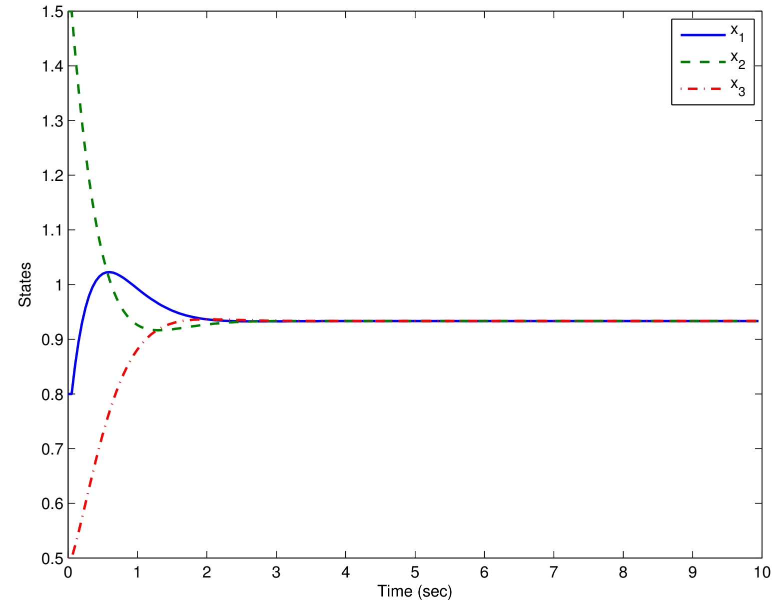

Figure 14: System responses

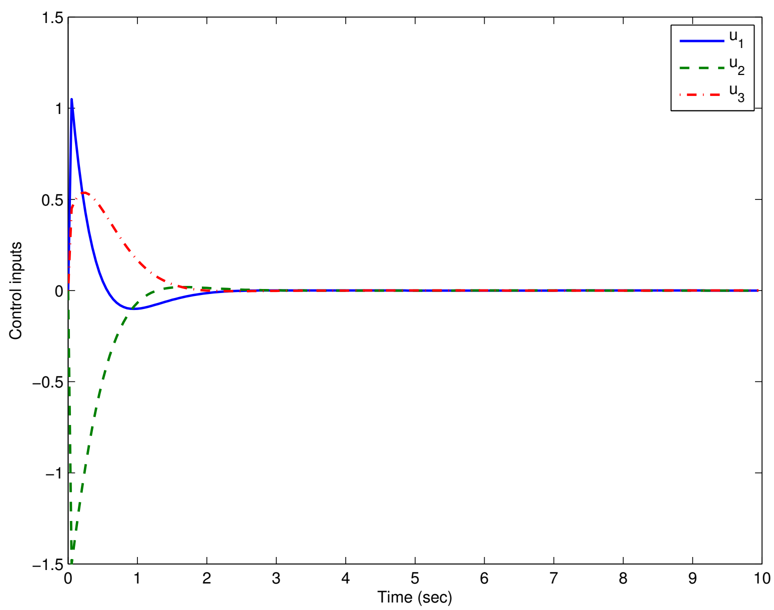

Figure 15: Cooperative control inputs

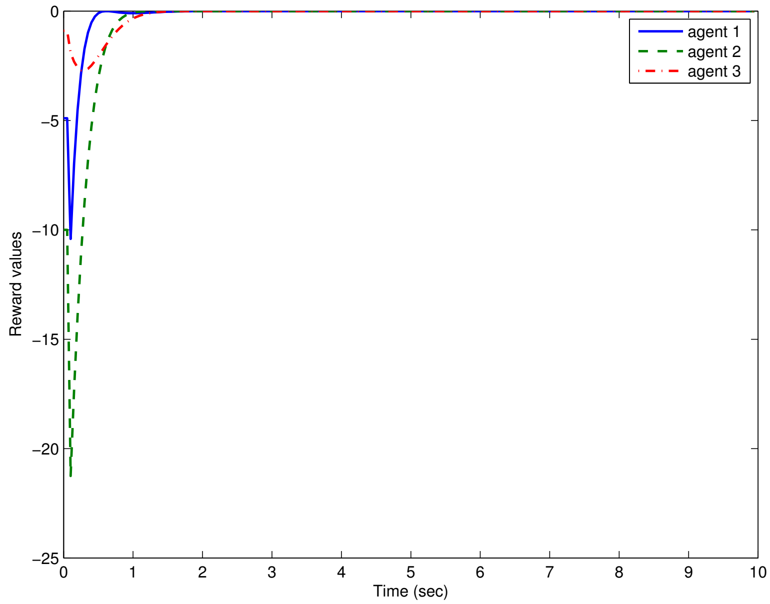

Figure 16: Performance Index

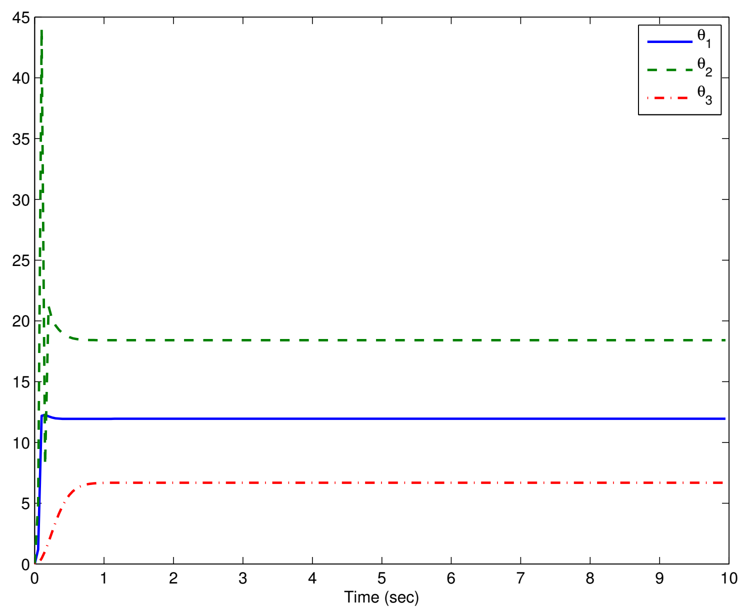

Figure 17: Parameters estimation

where a_i and g_i are constants, $x_i \in \Re$ state, and $u_i \in R$ control input. We first consider the case of a_i being unknown and g_i is known. For simplicity, we simply assume that $g_i = 1$. Let the reference trajectory be described by the first-order differential equation

$$\dot{x}_0 = a_0 x_0 + r_0(t) \tag{43}$$

where unknown constant $a_0 < 0$, $r_0(t)$ is a piecewise-continuous bounded function of time parameterized by $r_0(t) = \phi^T(t)w$, where basis functions $\phi(t) = [\phi_1(t), \phi_2(t), \cdots, \phi_l(t)]^T \in \Re^l$ are available to all agents, and parameters $w = [w_1, w_2, \cdots, w_l]^T \in \Re^l$ are unknown constants. The control objective is to design adaptive control u_i such that

$$\lim_{t\to\infty} x_i(t) = x_0(t).$$

5.3.2 Proposed Adaptive Control

To proceed, let us define the Laplacian matrix L for the sensing/communication $S(t)$ as follows.

$$L = \text{diag}\left\{\sum_{j=1}^{n} s_{ij}\right\} - S(t)$$

We also assume that the reference state $x_0(t)$ is available to at least one agent through sensing/communication detection, and this is described by a diagonal matrix B given below

$$B = \text{diag}\{b_{i0}\}$$

where $b_{i0} > 0$ means that agent i has the information $x_0(t)$. Let $\hat{a}_i$ be the parameter estimate of $a_i^* = a_0 - a_i$ for agent i, $\hat{w}_{ij}$ be the estimate of w_j by agent i, $\hat{w}_i = [\hat{w}_{i1}, \cdots, \hat{w}_{il}]^T$. The control input for agent i is chosen to be

$$u_i = \hat{a}_i x_i + \phi^T(t)\hat{w}_i \tag{44}$$

Defining the tracking error $\tilde{X} = [\tilde{x}_1, \cdots, \tilde{x}_n]^T = [x_1 - x_0, \cdots, x_n - x_0]^T$, the parameter errors $\tilde{a}_i = \hat{a}_i - a_i^*$, $\tilde{w}_i = \hat{w}_i - w_i = [\tilde{w}_{i1}, \cdots, \tilde{w}_{il}]^T$, we obtain the error equation as

$$\dot{\tilde{x}}_i = a_0\tilde{x}_i + \tilde{a}_i x_i + \phi^T \tilde{w}_i \tag{45}$$

and the overall error dynamics as

$$\dot{\tilde{X}} = a_0\tilde{X} + \mathcal{X}\tilde{a} + \sum_{i=1}^{l} \Phi_i \tilde{w}_{*i} \tag{46}$$

where $\tilde{a} = [\tilde{a}_1, \cdots, \tilde{a}_n]^T$, $\tilde{w}_{*i} = [\tilde{w}_{1i}, \cdots, \tilde{w}_{ni}]^T$, $\mathcal{X} = \text{diag}[x_1, \cdots, x_n]$, and $\Phi_1 = \text{diag}[\phi_i(t), \cdots, \phi_i(t), \cdots, \phi_i(t)]$.

We further define consensus error $e_i = \sum_{j\in\mathcal{N}_i} s_{ij}(x_j - x_i) + b_{i0}(x_0 - x_i)$. It thus can be verified that

$$(L + B)\tilde{X} = [e_1, \cdots, e_i, \cdots, e_n]^T \tag{47}$$

Let us consider the Lyapunov-like function

$$V = \frac{1}{2}\tilde{X}^T(L+B)\tilde{X} + \frac{1}{2}\sum_{i=1}^{n}\Gamma_{a_i}(\tilde{a}_i)^2 + \frac{1}{2}\sum_{j=1}^{l}\sum_{i=1}^{n}\Gamma_{w_{ij}}(\tilde{w}_{ij})^2 \tag{48}$$

where $\Gamma_{a_i} > 0$, $\Gamma_{w_{ij}} > 0$. The time derivative of V along the trajectories of (46) is given by

$$\dot{V} = \tilde{X}^T(L+B)\dot{\tilde{X}} + \sum_{i=1}^{n}\Gamma_{a_i}\tilde{a}_i\dot{\tilde{a}}_i + \sum_{j=1}^{l}\sum_{i=1}^{n}\Gamma_{w_{ij}}\tilde{w}_{ij}\dot{\tilde{w}}_{ij} \tag{49}$$

Choosing the adaptive laws

$$\dot{\hat{a}}_i = \Gamma_{a_i}^{-1}x_{i1}e_i \tag{50}$$

$$\dot{\hat{w}}_{ij} = \Gamma_{w_{ij}}^{-1}\phi_j e_i \tag{51}$$

for $i = 1, \cdots, n$, $j = 1, \cdots, l$, and then we obtain

$$\dot{V} = a_0\tilde{X}^T(L+B)\tilde{X} \leq 0 \tag{52}$$

which implies $\tilde{X}$, $\hat{a}_i$, $\hat{w}_{ij} \in \mathcal{L}_\infty$. Also $\tilde{X} \in \mathcal{L}_2$ and $\dot{X} \in \mathcal{L}_\infty$, which further implies that $\tilde{X} \to 0$ as $t \to \infty$. The main result is summarized into the following theorem.

Theorem 4 *Consider the multiagent system in (42). If the sensing/communication topology $S(t)$ is connected, and B has at least one entry being nonzero, then the distributed adaptive cooperative control in (44) together with the adaptive laws in (50) and (51) guarantee the boundedness of all signals of the closed-loop system and achieve asymptotical consensus tracking.*

5.3.3 Simulation Results

The proposed adaptive cooperative control is simulated for the following multiagent system with three agents

$$\dot{x}_1 = 2x_1 + u_1$$

$$\dot{x}_2 = x_2 + u_2$$

$$\dot{x}_3 = 5x_3 + u_3$$

Assume the Informed Agent: $\dot{x}_0 = -x_0 + r(t)$, $r(t) = 2\cos(t) + 3\cos(2t)$. The sensing/communication matrix S and leader information matrix B are given by

$$S = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

The corresponding adaptive consensus control is

$$u_i = \hat{a}_i x_i + \hat{w}_{i1}\cos(t) + \hat{w}_{i2}\cos(2t)$$

withe adaptive laws

$$\dot{\hat{a}}_i = kx_i(\sum_j s_{ij}(x_j - x_i) + b_{i0}(x_0 - x_i))$$

$$\dot{\hat{w}}_{i1} = k\cos(t)(\sum_j s_{ij}(x_j - x_i) + b_{i0}(x_0 - x_i))$$

$$\dot{\hat{w}}_{i2} = k\cos(2t)(\sum_j s_{ij}(x_j - x_i) + b_{i0}(x_0 - x_i))$$

Simulation results are given in figures 18-23, which illustrate the effectiveness of the proposed design.

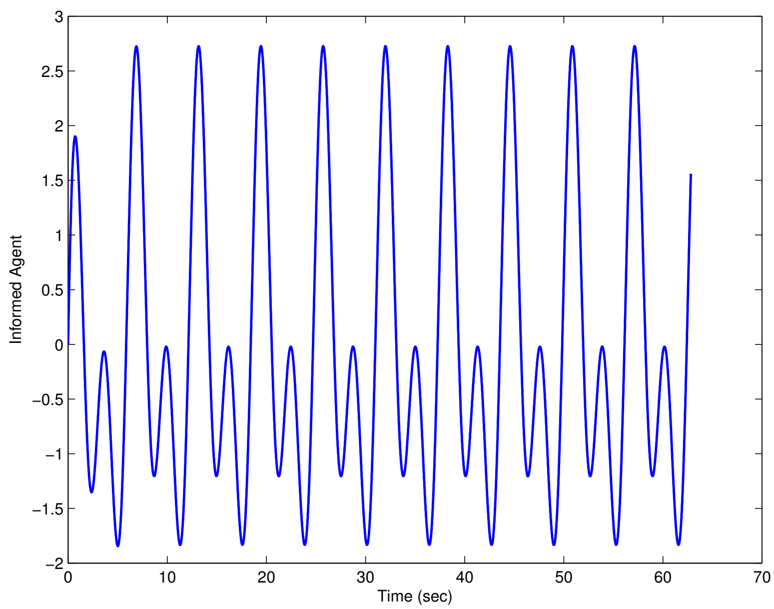

Figure 18: The trajectory of informed agent

6.0 RESULTS AND DISCUSSION: MULTIAGENT COORDINATION APPLICATIONS

In this section, we present two case studies for multiagent coordination tasks. One is for formation control of mobile robots, and the other is for coverage control of mobile agents.

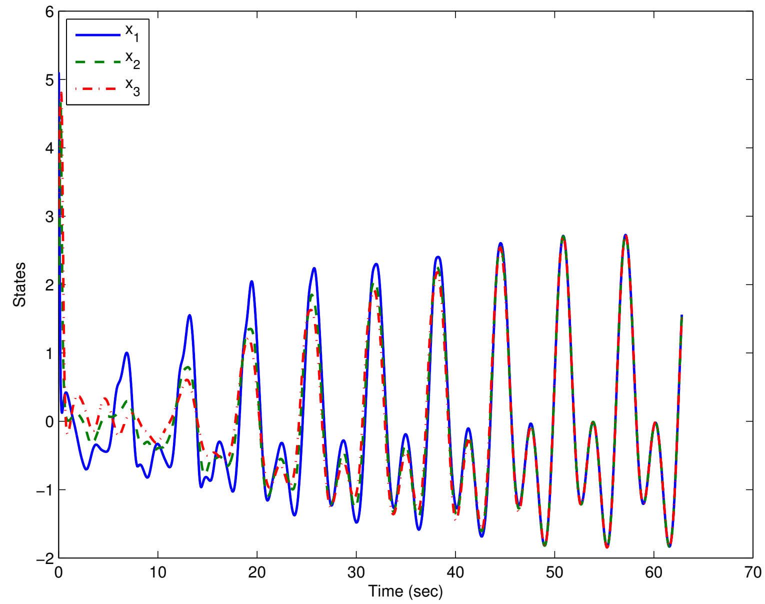

Figure 19: The trajectories of agents

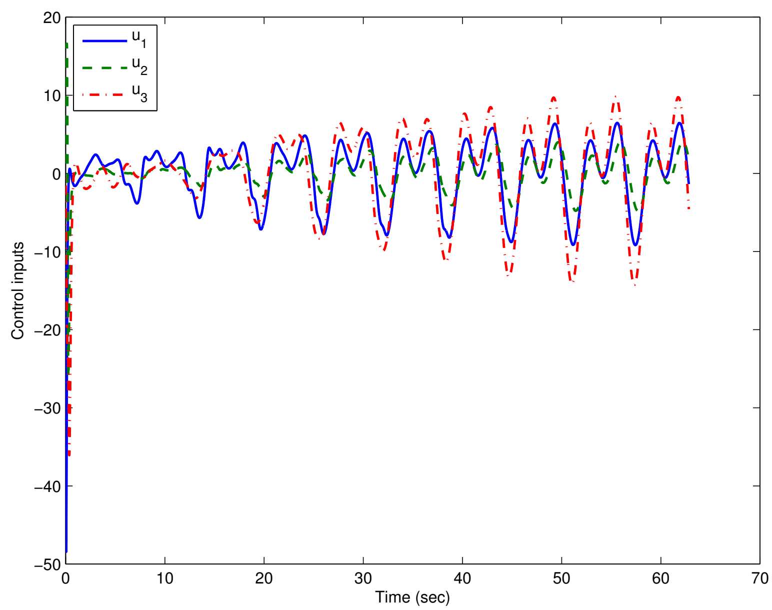

Figure 20: Adaptive cooperative controls

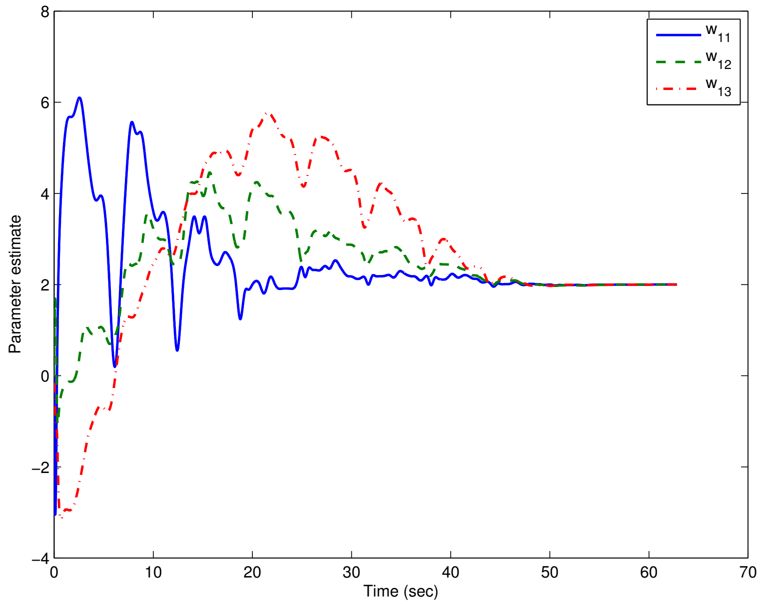

Figure 21: Parameter estimate for w_1

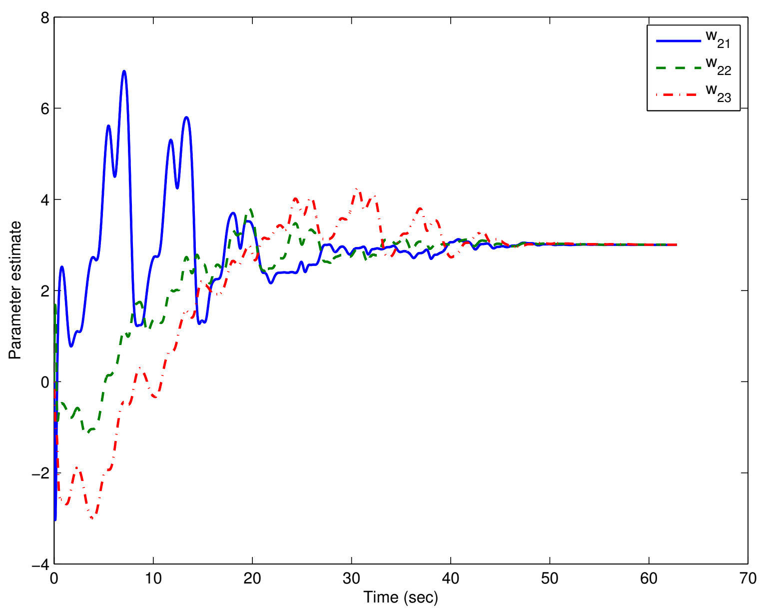

Figure 22: Parameter estimate for w_2

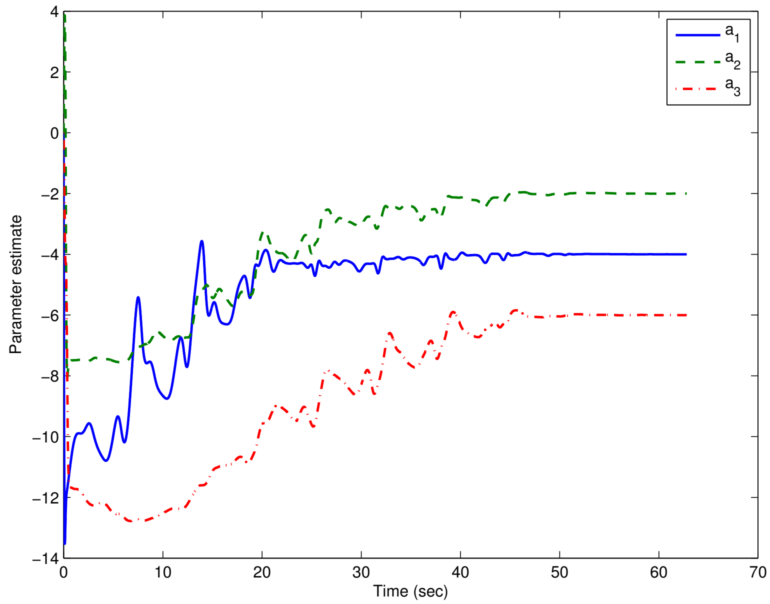

Figure 23: Parameter estimate for a

6.1 Multiagent Formation Control

6.1.1 Problem Statement

Consider a network of multiple nonholonomic mobile robots with the individual system dynamics given by

$$\begin{aligned} \dot{x}_i &= v_i \cos\theta_i, \\ \dot{y}_i &= v_i \sin\theta_i, \\ \dot{\theta}_i &= \omega_i \end{aligned} \tag{53}$$

where $i \in \Omega \triangleq \{1, \cdots, n\}$, $(x_i, y_i) \in \Re^2$ denotes the ith robot's position, θ_i is the orientation, $v_i \in \Re$ driving velocity, and $\omega_i \in \Re$ the steering velocity.

The design objective of this paper is to coordinate the motion of individual robots to follow a desired trajectory contour while maintaining certain prescribed geometric formation shape through local information exchange among robots. By taking the whole group of mobile robots as a virtual body moving along the desired trajectory, formation shape of robots in the group can be determined by a set of local coordinates with reference to the moving frame attached to the desired trajectory.

More specifically, let $q_0(t) = [x_0(t), y_0(t)]^T \in \Re^2$ be the desired trajectory for the group motion, the moving frame $\mathcal{F}(t)$ attached to $q_0(t)$ can be defined by the following orthonormal vectors $e_1(t)$

and $e_2(t)$

$$e_1(t) = \begin{bmatrix} e_{11}(t) \\ e_{12}(t) \end{bmatrix} = \begin{bmatrix} \dfrac{\dot{x}_0(t)}{\sqrt{[\dot{x}_0(t)]^2 + [\dot{y}_0(t)]^2}} \\ \dfrac{\dot{y}_0(t)}{\sqrt{[\dot{x}_0(t)]^2 + [\dot{y}_0(t)]^2}} \end{bmatrix},$$

$$e_2(t) = \begin{bmatrix} e_{21}(t) \\ e_{22}(t) \end{bmatrix} = \begin{bmatrix} -\dfrac{\dot{y}_0(t)}{\sqrt{[\dot{x}_0(t)]^2 + [\dot{y}_0(t)]^2}} \\ \dfrac{\dot{x}_0(t)}{\sqrt{[\dot{x}_0(t)]^2 + [\dot{y}_0(t)]^2}} \end{bmatrix}.$$

Accordingly, any formation consisting of n robot positions in $\mathcal{F}(t)$ can be expressed as $\{P_1, \cdots, P_n\}$ with

$$P_i(t) = \alpha_{i1}e_1(t) + \alpha_{i2}e_2(t), \tag{54}$$

where α_{ij} are constants of determining the formation. To this end, the formation control objective can be recast as to design the control laws $v_i(t)$ and $\omega_i(t)$ for the ith robot such that

$$\lim_{t\to\infty} \left[\begin{bmatrix} x_i(t) \\ y_i(t) \end{bmatrix} - q_0(t) - P_i(t) \right] = 0. \tag{55}$$

6.1.2 Proposed Linearization-Based Control

To facilitate the control design, the robot model (71) is first converted into a linear model as

$$\dot{z}_{i1} = u_{i1}, \quad \dot{z}_{i2} = u_{i2} \tag{56}$$

where $z_{i1} = x_i + R\cos\theta_i$, $z_{i2} = y_i + R\sin\theta_i$, and

$$\begin{bmatrix} v_i \\ \omega_i \end{bmatrix} = \begin{bmatrix} \cos\theta_i & \sin\theta_i \\ -\frac{\sin\theta_i}{R} & \frac{\cos\theta_i}{R} \end{bmatrix} \begin{bmatrix} u_{i1} \\ u_{i2} \end{bmatrix}$$

To this end, the proposal new coordination control is of the form

$$u_{i1} = \sum_{j\in\mathcal{N}_i} \alpha_{ij}(t)\mathrm{sgn}\left(z_{j1} - z_{i1} + \sum_{l=1}^{2}(a_{il} - a_{jl})e_l \right) + \dot{q}_0 + \dot{p}_i^d \tag{57}$$

$$u_{i2} = \sum_{j\in\mathcal{N}_i} \alpha_{ij}(t)\mathrm{sgn}\left(z_{j2} - z_{i2} + \sum_{l=1}^{2}(a_{il} - a_{jl})e_l \right) + \dot{q}_0 + \dot{p}_i^d \tag{58}$$

in which the control gains are chosen based on the following guidelines: for any agent l,

1) if $z_l(t_k^s) = \max_{j\in\mathcal{N}_l} z_j(t_k^s) = \min_{j\in\mathcal{N}_l} z_j(t_k^s)$, then $\alpha_{lj}(t_k^s)$ can be any bounded positive value.

2) if $z_l(t_k^s) \geq \max_{j\in\mathcal{N}_l} x_j(t_k^s)$, let $\alpha_{lj}(t_k^s)$ be selected to satisfy the inequality

$$0 \leq \sum_{j\in\mathcal{N}_l} \alpha_{lj}(t_k^s) < \frac{z_l(t_k^s) - \min_{j\in\mathcal{N}_l} z_j(t_k^s)}{\overline{c}_t}$$

3) if $z_l(t_k^s) \leq \min_{j\in\mathcal{N}_l} z_j(t_k^s)$, let $\alpha_{lj}(t_k^s)$ be selected to satisfy the inequality

$$0 \leq \sum_{j\in\mathcal{N}_l} \alpha_{lj}(t_k^s) < \frac{\max_{j\in\mathcal{N}_l} z_j(t_k^s) - z_l(t_k^s)}{\overline{c}_t}$$

4) if $\min_{j\in\mathcal{N}_l} z_j(t_k^s) < z_l(t_k^s) < \max_{j\in\mathcal{N}_l} z_j(t_k^s)$, let $\alpha_{lj}(t_k^s)$ be selected to satisfy the inequality

$$0 \leq \sum_{j\in\mathcal{N}_l} \alpha_{lj}(t_k^s) < \min\left(\begin{array}{c} \frac{\max_{j\in\mathcal{N}_l} z_j(t_k^s) - z_l(t_k^s)}{\overline{c}_t}, \\ \frac{z_l(t_k^s) - \min_{j\in\mathcal{N}_l} z_j(t_k^s)}{\overline{c}_t} \end{array} \right)$$

6.1.3 Simulation and Experimental Results

We conducted the experimental study to verify the proposed design. In the experiment, we use 4 Amigobot robots (see figure 24) and the sensing/communication topology among robots is assumed to be

$$S = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

The testing results validated the proposed design.

Figures 25 and 26 show the snapshots of experiments conducted using AmigoBots at Robotics Lab, Bethune-Cookman University.

Figure 26 shows the formation shape changing from rectangle to line, then to rhombus, and finally converging to one point.

6.1.4 Proposed Nonlinear Model-Based Control with Limited Information of Desired Trajectory

It is noted that the control objective defined in (55) can be achieved through the standard tracking control design for individual robots if the desired trajectory $q_0(t)$ and its derivative $\dot{q}_0(t)$ are available to every robot. However, such a design may not be robust in the presence of disturbance and noise measurements due to the lack of coordination among robots. On the other hand, the desired trajectory $q_0(t)$ may be known only by some of robots in the group. Therefore, it is desirable to design distributed formation control law for the ith robot based on information exchange and relative position measurement between robots within its sensing/communication range.

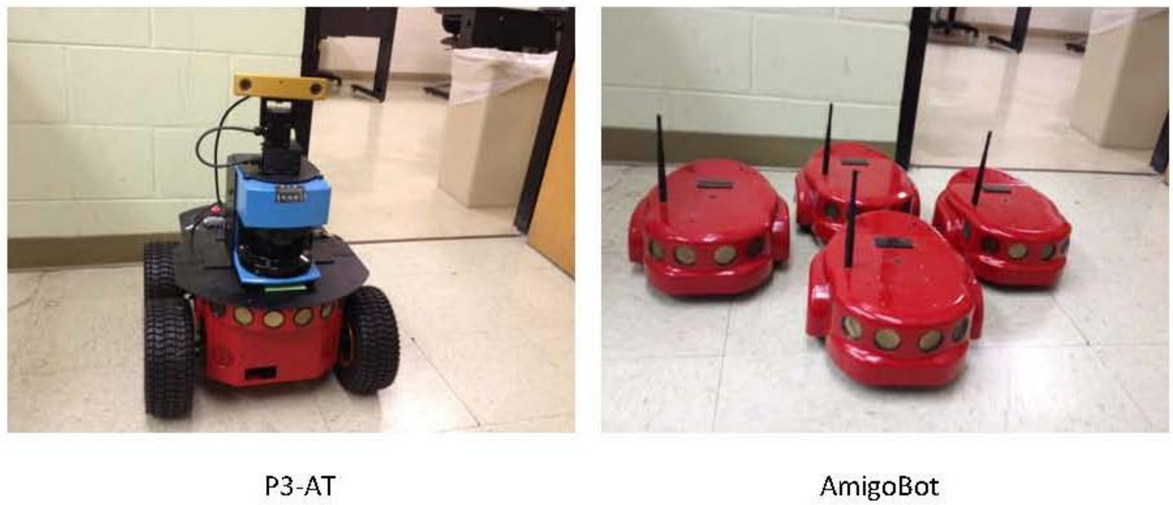

Figure 24: P3-AT and AmigoBots

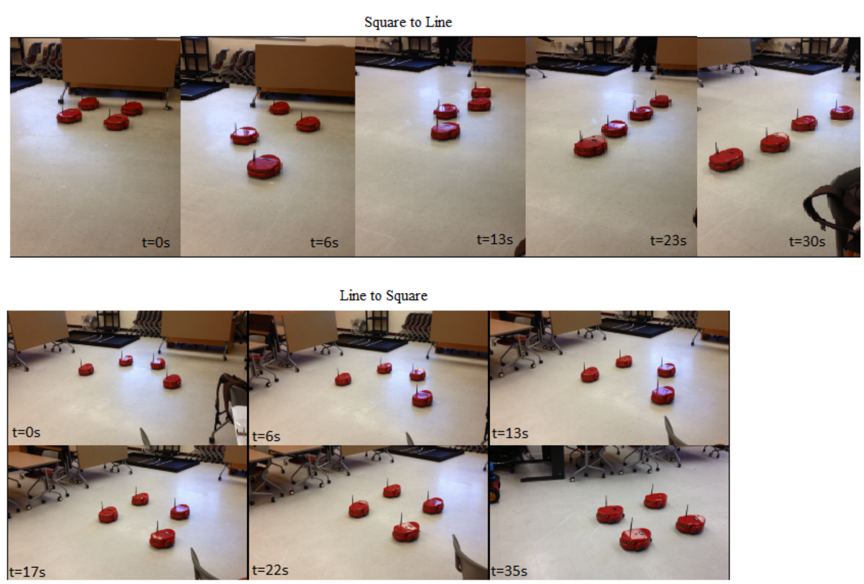

Figure 25: Rectangle-to-line and line-to-rectangle formation control with undirected communication

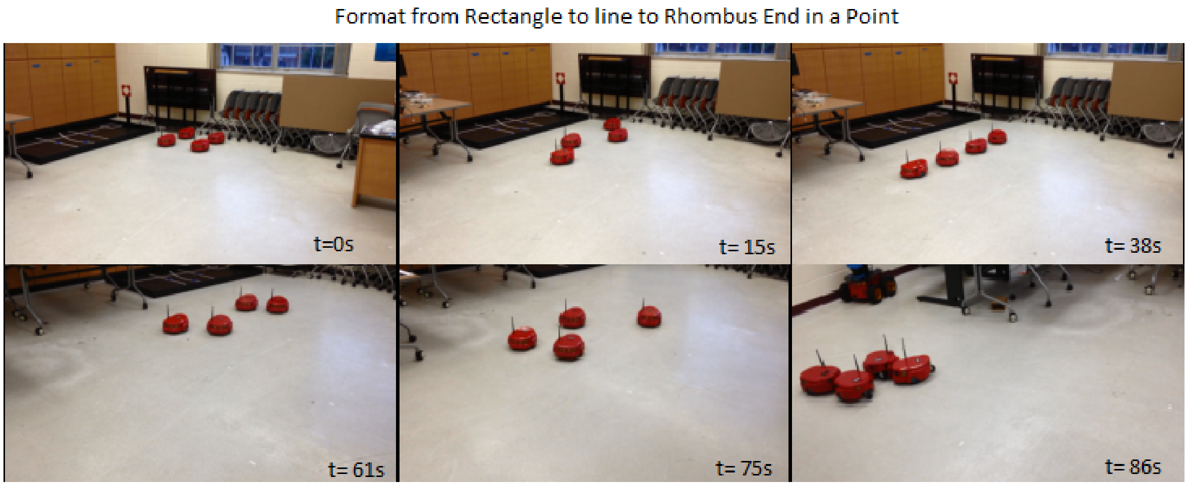

Figure 26: Formation changes from rectangle to line to rhombus and ends with a point with undirected communication

Distributed Observers for Desired Trajectory

The proposed new formation control is done with the aid of distributed observes for the estimation of $q_0(t)$. The proposed distributed observer is of the form (for $t \in [t_k^s, t_{k+1}^s)$)

$$\dot{x}_{i,0}(t) = \sum_{j \in \mathcal{N}_i} \alpha_{ij} \text{sgn}(x_{j,0}(t_k^s) - x_{i,0}(t_k^s)) + \alpha_{i0} s_{i0} \text{sgn}(x_0(t_k^s) - x_{i,0}(t_k^s)) \tag{59}$$

$$\dot{y}_{i,0}(t) = \sum_{j \in \mathcal{N}_i} \alpha_{ij} \text{sgn}(y_{j,0}(t_k^s) - y_{i,0}(t_k^s)) + \alpha_{i0} s_{i0} \text{sgn}(y_0(t_k^s) - y_{i,0}(t_k^s)) \tag{60}$$

$$\dot{\theta}_{i,0}(t) = \sum_{j \in \mathcal{N}_i} \alpha_{ij} \text{sgn}(\theta_{j,0}(t_k^s) - \theta_{i,0}(t_k^s)) + \alpha_{i0} s_{i0} \text{sgn}(\theta_0(t_k^s) - \theta_{i,0}(t_k^s)) \tag{61}$$

where $x_{i,0}(t)$, $y_{i,0}(t)$ and $\theta_{i,0}(t)$ are the ith robot's estimate of $x_0(t)$, $y_0(t)$, and $\theta_0(t)$, respectively, $s_{i0} = 1$ if and only if the ith robot has the direct access to the information of the desired trajectory, $\alpha_{i,j}$ and α_{i0} are piecewise constant control gains to be designed, and sgn$(\cdot)$ function is defined as

$$\text{sgn}(z) = \begin{cases} 1, & z > 0 \\ 0, & z = 0 \\ -1, & z < 0 \end{cases}$$

Nonlinear Formation Control with Limited Information of a Desired Trajectory

The proposed new design is summarized into the following proposition.

Theorem 5 *Consider a group of nonholonomic mobile robots. Let the distributed cooperative control be for* $t \in [t_0 + kT_s, t_0 + (k+1)T_s)$

$$
\begin{aligned}
u_{i1}(t) &= a_{i1}^k + a_{i2}^k \sin \omega(t - t_0 - kT_s) && (62)\\
u_{i2}(t) &= b_{i1}^k + b_{i2}^k \cos \omega(t - t_0 - kT_s) && (63)
\end{aligned}
$$

where $\omega = \frac{2\pi}{T_s}$, $a_{i2}^k \neq 0$ *can be any constant, and*

$$
\begin{aligned}
a_{i1}^k &= \frac{1}{T_s} \sum_{j=1}^n G_{ij}(k)[x_j(k) - x_i(k) - x_j^d(k) + x_i^d(k+1)], && (64)\\
b_{i1}^k &= \frac{1}{T_s} \sum_{j=1}^n G_{ij}(k)[z_{j2}(k) - z_{i2}(k)], && (65)\\
b_{i2}^k &= \frac{2\omega}{a_{i2}^k T_s} \left[\sum_{j=1}^n G_{ij}(k)[y_j(k) - y_i(k) - y_j^d(k) + y_i^d(k+1)] - \frac{a_{i1}^k b_{i1}^k T_s^2}{2} \right.\\
&\quad \left. - a_{i1}^k z_{i2}(k) T_s + \frac{a_{i2}^k b_{i1}^k T_s}{\omega} \right]. && (66)
\end{aligned}
$$

with

$$
\begin{aligned}
G_{ij}(k) &= \frac{s_{ij}(k)}{\sum_{\eta=1}^n s_{i\eta}(k)}, \quad j = 1, \cdots, n. && (67)\\
x_i^d(k) &= x_{i,0}(k) + \alpha_{i1} \cos \theta_{i,0}(k) - \alpha_{i2} \sin \theta_{i,0}(k)\\
y_i^d(k) &= y_{i,0}(k) + \alpha_{i1} \sin \theta_{i,0}(k) + \alpha_{i2} \cos \theta_{i,0}(k)
\end{aligned}
$$

$$
\begin{aligned}
x_i^d(k+1) &= x_{i,0}(k+1) + \alpha_{i1} \cos \theta_{i,0}(k+1) - \alpha_{i2} \sin \theta_{i,0}(k+1)\\
y_i^d(k+1) &= y_{i,0}(k+1) + \alpha_{i1} \sin \theta_{i,0}(k+1) + \alpha_{i2} \cos \theta_{i,0}(k+1)
\end{aligned}
$$

where $x_{i,0}$, $y_{i,0}$, *and* $\theta_{i,0}$ *are governed by (59), (60), and (61), respectively, and*

$$
\begin{aligned}
x_{i,0}(k+1) &= x_{i,0}(k) + T_s(\sum_{j \in \mathcal{N}_i} \alpha_{ij} sgn(x_{j,0}(k) - x_{i,0}(k)) + \alpha_{i0} s_{i0} sgn(x_0(k) - x_{i,0}(k))) && (68)\\
y_{i,0}(k+1) &= y_{i,0}(k) + T_s(\sum_{j \in \mathcal{N}_i} \alpha_{ij} sgn(y_{j,0}(k) - y_{i,0}(k)) + \alpha_{i0} s_{i0} sgn(y_0(k) - y_{i,0}(k))) && (69)\\
\theta_{i,0}(k+1) &= \theta_{i,0}(k) + T_s(\sum_{j \in \mathcal{N}_i} \alpha_{ij} sgn(\theta_{j,0}(k) - \theta_{i,0}(k)) + \alpha_{i0} s_{i0} sgn(\theta_0(k) - \theta_{i,0}(k))) && (70)
\end{aligned}
$$

Then the formation control objective (55) is achieved.

6.1.5 Simulation Results

We simulate the proposed formation control by considering three mobile robots moving according to a circular contour while maintaining a right triangle formation.

Let $q_0(t)$ be $[\sin(0.2t), -\cos(0.2t)]^T$. The corresponding moving frame is given by $e_1(t) = [\cos(0.2t), -\sin(0.2t)]^T$, $e_2(t) = [\sin(0.2t), \cos(0.2t)]^T$. The formation parameters are given by $\alpha_{11} = 0, \alpha_{12} = 0, \alpha_{21} = -1, \alpha_{22} = 1, \alpha_{31} = -1, \alpha_{32} = -1$. The initial conditions $[x_i(t_0), y_i(t_0), \theta_i(t_0)]$ are given by $[0.1, 0.2, \pi/4], [1, -2, \pi/6], [-1, -1.5, 0]^T$ for $i = 1, 2, 3$, respectively, $a_{i2}^k = 0.2$ and $T_s = 0.1$. Figure 27 illustrates the phase portrait under the proposed controls proposed controls (62) and (63).

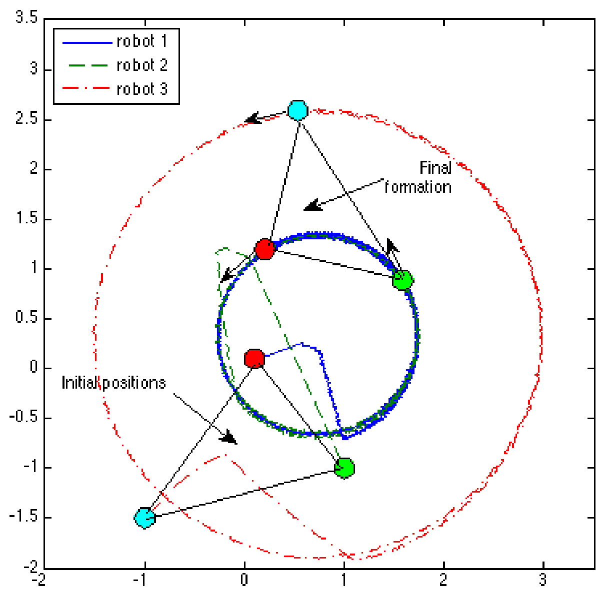

Figure 27: Phase portrait of three robots

6.2 Multiagent Coverage Control

Coverage control aims to address the issue of deployment of sensor networks for tasks like monitoring an environment, environment modeling, search and rescue, and so on. In recent years, mobile autonomous agents have been applied in the construction of mobile sensor networks due to their flexibility and resilience to dynamically changing environments.

In this section, we propose a distributed deployment algorithm for a group of mobile robots to cover a convex region. The individual mobile robot considered has kinematic constraints,

and may only exchange information locally with its neighboring counterparts due to its limited sensing/communication range. The proposed deployment algorithm iteratively updates the Voronoi partition through local information exchange, and then moves toward its centroid based on centroid-driven control algorithms.

6.2.1 Problem Statement

To solve the autonomous deployment problem, we make the following assumptions without loss of generality:

- The robots have the knowledge of the area to be covered and sensed.
- The robots have limited sensing ranges r_s, and limited communication ranges r_c. That is, only points in a circle centered at the current robot's position and of radius r_s can be sensed by the robot. In addition, at time t, robot i can communicate with its neighboring robot j, $j \in \mathcal{N}_i(t) = \{j | d_{ij} \leq r_c\}$, where d_{ij} is the distance between the ith robot and the jth robot.
- For a given region, there are enough number of n mobile robotic agents to completely cover the area.

To this end, the multiagent coverage control problem is formulated as designing a distributed deployment control algorithm to move the robots towards the centroid of the corresponding partitioned regions based on the minimization of certain coverage cost functions. Under the aforementioned assumptions, the coverage control problem has at least one solution. In this section, a new paradigm is proposed to deploy the robots by assuming limited sensing and communication ranges.

6.2.2 Proposed Coverage Control

The kinematic model of mobile robotic agent carrying sensors is described by the following equations:

$$\begin{aligned} \dot{x}_i &= v_i \cos\theta_i, \\ \dot{y}_i &= v_i \sin\theta_i, \\ \dot{\theta}_i &= \omega_i \end{aligned} \tag{71}$$

where $i \in \Omega \stackrel{\triangle}{=} \{1, \cdots, n\}$, $p_i = [x_i, y_i]^T \in \Re^2$ denotes the ith robot's position, θ_i is the orientation, $v_i \in \Re$ driving velocity, and $\omega_i \in \Re$ the steering velocity. The optimal coverage control problem is then defined as designing distributed cooperative control v_i and ω_i such that agents converge to optimal positions p_i^* by minimizing certain cost function.

The proposed deployment algorithm is a recursive one. At each sampling time instant, each robot first computes its Voronoi cell based on its communication with neighboring robots, then determine the centroid of its Voronoi region, and then moves towards it by employing a distributed coordination algorithm.

In what follows, we describe the basic idea of Voronoi partition based coverage control for mobile robots. Let us denote an arbitrary point in the region Q as q. At each sampling time instant, the agents will be able to generate the Voronoi partition of Q. That is, for agent i at position p_i, its Voronoi region satisfies

$$V_i = \{q \in Q | \|q - p_i\| \leq \|q - p_j\|, \forall j \neq i\} \tag{72}$$

Define cost function over the region as

$$J(p_1, \cdots, p_n) = \sum_i^n \int_{V_i} \frac{1}{2} \|q - p_i\|^2 \phi(q) dq \tag{73}$$

where $\phi(q)$ is a weighting function of importance over Q. The distance function $\frac{1}{2}\|q - p_i\|^2$ is included in the cost function for the consideration of reducing energy consumed by a sensor's transceiver because it is generally a function of distance. In addition, the reliability of the information at q measured by robot at p_i will degrade with the increase of distance $\|q - p_i\|^2$.

At each sampling time instant, the generation of Voronoi region V_i for robot i is based on the robots in its neighboring set $\mathcal{N}_i$. That is, robot i can only use the position information of the robots in its communication range r_c to compute V_i. This is a realistic situation since during the motion, the robot could move in or out the communication range which is limited.

Once the Voronoi region is obtained, a simple control to drive the robot to the centroid of the Voronoi region is to follow negative gradient of cost function J, that is,

$$-\frac{\partial J}{\partial p_i} = -\int_{V_i} (q - p_i)\phi(q) dq$$

However, as discussed before, the kinematic model in (71) is nonlinear and may not be able to follow the negative gradient due to velocity constraints. A simple way to avoid this issue is to conduct input/output linearization by choosing a reference point off the robot center (x_i, y_i), that is, let the cartesian coordinates of the off-center reference point be

$$\begin{aligned} p_{i1} &= x_i + b\cos\theta_i & (74) \\ p_{i2} &= y_i + b\sin\theta_i & (75) \end{aligned}$$

where $b > 0$ is a constant. Differentiating (74) and (75) with respect to time, we have

$$\begin{aligned} \begin{bmatrix} \dot{p}_{i1} \\ \dot{p}_{i2} \end{bmatrix} &= \begin{bmatrix} \cos\theta & -b\sin\theta \\ \sin\theta & b\cos\theta \end{bmatrix} \begin{bmatrix} v_i \\ \omega_i \end{bmatrix} \\ &\triangleq T(\theta) \begin{bmatrix} v_i \\ \omega_i \end{bmatrix} \triangleq \begin{bmatrix} u_{i1} \\ u_{i2} \end{bmatrix} & (76) \end{aligned}$$

To this end, the distributed deployment control for robot i is given by

$$u_i \triangleq \begin{bmatrix} u_{i1} \\ u_{i2} \end{bmatrix} = -\frac{\partial J}{\partial p_i} = -\int_{V_i} (q - p_i)\phi(q) dq = -M_{V_i}(C_{V_i} - p_i) \tag{77}$$

where mass M_{V_i} is given by

$$M_{V_i} = \int_{V_i} \phi(q) dq \tag{78}$$

the first moment

$$L_{V_i} = \int_{V_i} q\phi(q) dq \tag{79}$$

and the centroid

$$C_{V_i} = \frac{L_{V_i}}{M_{V_i}} \tag{80}$$

Once u_i is obtained, the control inputs v_i and ω_i can be calculated by using inverse input transformation given below:

$$\begin{bmatrix} v_i \\ \omega_i \end{bmatrix} = T^{-1}(\theta) \begin{bmatrix} u_{i1} \\ u_{i2} \end{bmatrix}.$$

6.2.3 Simulation Results

In this section, we simulate the proposed distributed deployment algorithm. Consider first the case with 5 mobile robotic agents, and we assume fully connected communication topology. That is, at each time instant, each robot has the position information of the rest members in the group. Figure 28 and 29 illustrate the initial location with Voronoi partition and the final position with Voronoi partition, respectively. Figure 30 illustrates of the evolution of the robots.

In the 2nd case, we consider 10 robots with limited communication ranges. Assume that the initial communication topology is defined by

$$C(0) = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

and changes subsequently based on system evolution. Figure 31 and 32 illustrate the initial location with Voronoi partition and the final position with Voronoi partition, respectively. Figure 33 illustrates of the evolution of the robots.

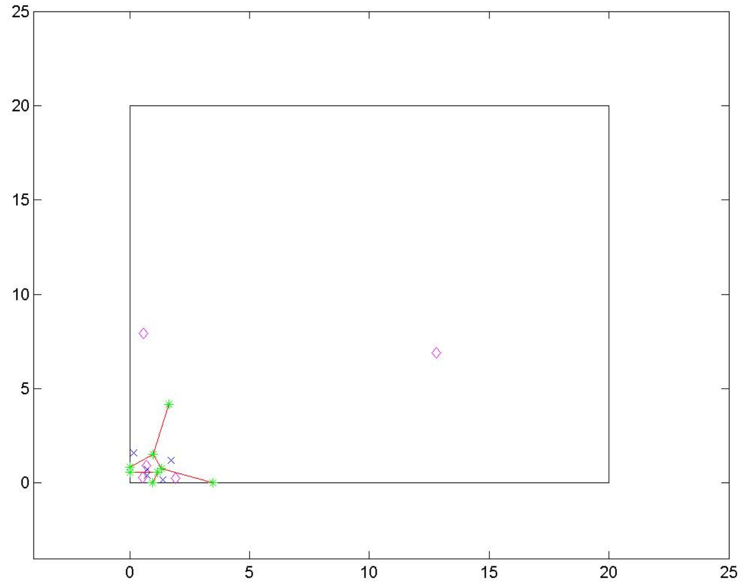

Figure 28: Initial location and Voronoi partition

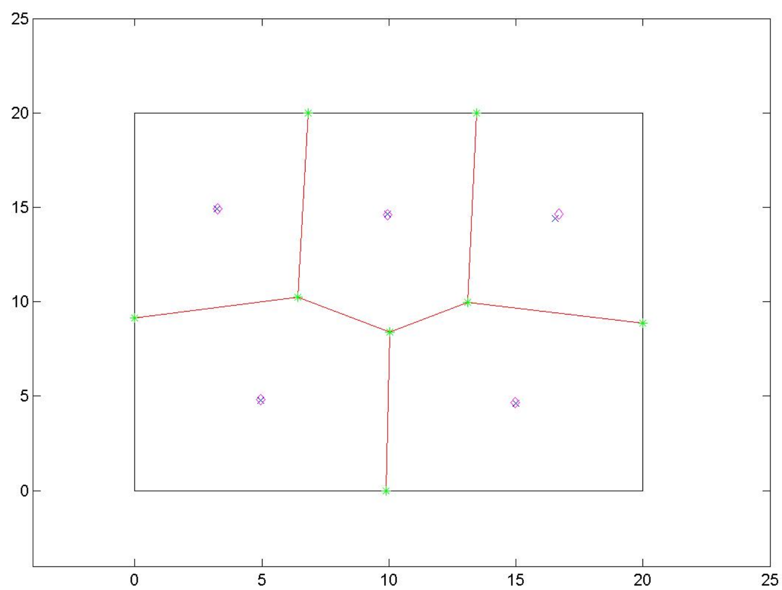

Figure 29: Final location and Voronoi partition

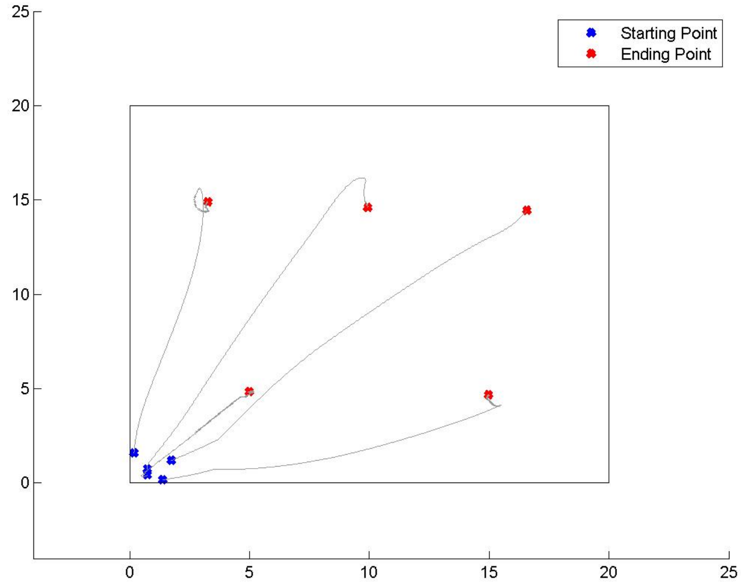

Figure 30: Evolution of the robots

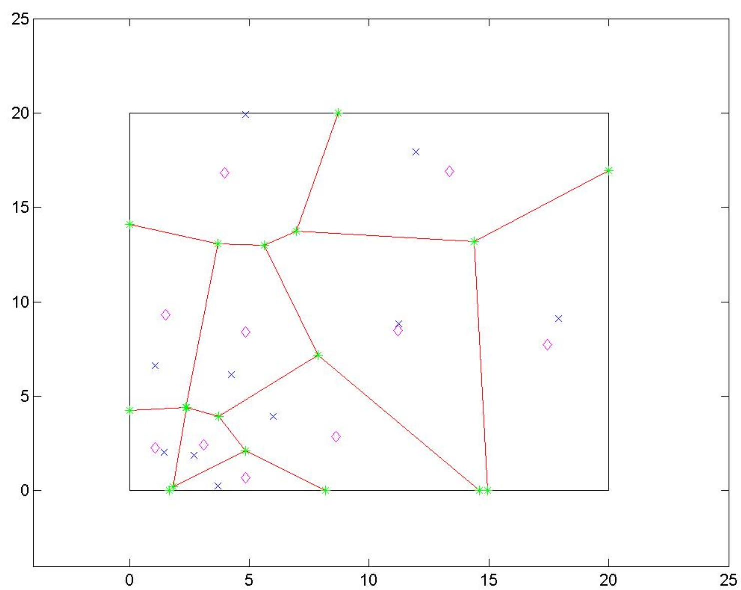

Figure 31: Initial location and Voronoi partition

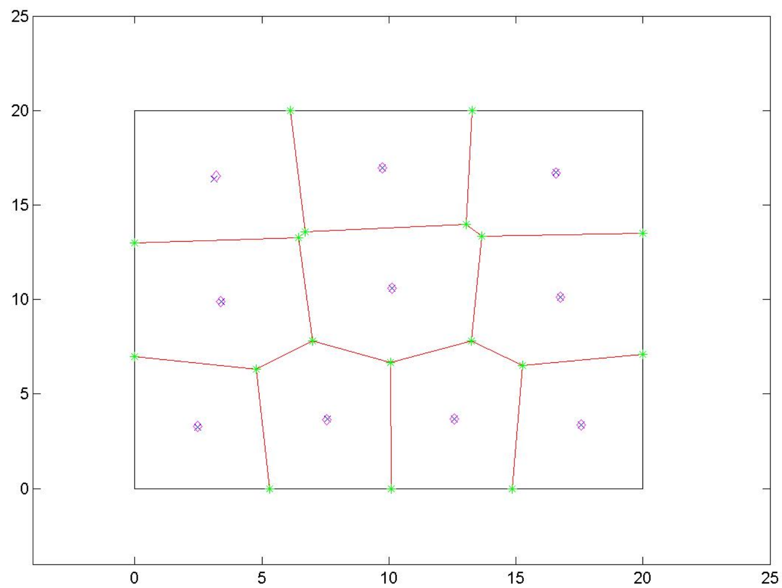

Figure 32: Final location and Voronoi partition

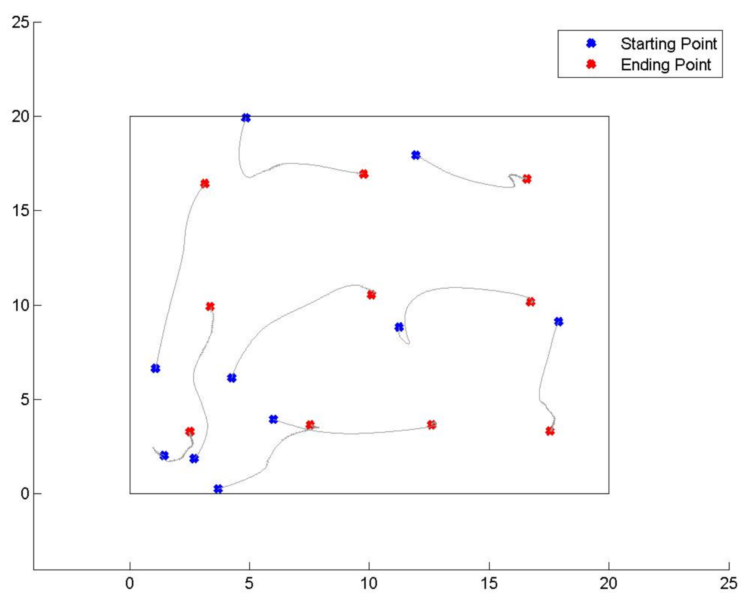

Figure 33: Evolution of the robots

7.0 CONCLUSIONS

In this report, we have presented a comprehensive description of the research results obtained through this project. Specifically, technique details have been provided for three sets of multiagent coordination algorithms which solve the distributed task coordination problem while ensuring system stability, accommodating the least-restrictive sensing/communication conditions, handling system uncertainties, and guaranteeing near-optimal performance. Computer simulation and experimental results are presented to illustrate the effectiveness of the proposed algorithms.

8.0 REFERENCES

[1] D. Bauso, L. Giarre, and R. Pesenti. Nonlinear protocols for optimal distributed consensus in networks of dynamic agents. *Systems and Control Letters*, 55:918–928, 2006.

[2] F. Bullo, J. Cortés, and S. Martínez. *Distributed Control of Robotic Networks*. Applied Mathematics Series. Princeton University Press, 2009. Electronically available at http://coordinationbook.info.

[3] Y. Cao and W. Ren. Optimal linear consensus algorithms: an lqr perspective. *IEEE Trans. on System, Man and Cybernetics Part B: Cybernetics*, 40:819–830, 2010.

[4] Y. Cao and W. Ren. Distributed coordinated tracking with reduced interaction via a variable structure approach. *IEEE Trans. Automa. Control*, 56:in press, 2011.

[5] D. V. Dimarogonas and K. J. Kyriakopoulos. Connectedness preserving distributed swarm aggregation for multiple kinematic robots. *IEEE Transactions on Robotics*, 24:1213–1223, 2008.

[6] J. A. Fax and R. M. Murray. Information flow and cooperative control of vehicle formations. *IEEE Transactions on Automatic Control*, 49:1465–1476, 2004.

[7] A. Jadbabaie, J. Lin, and A.S. Morse. Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Trans. on Automatic Control*, 48:988–1001, 2003.

[8] J.Shao, J. Wang, and T. Yang. Experimental validation of distributed cooperative control for mobile agents with switching topologies and time delays. In *IEEE International Conference on Electro/Information Technology*, Milwaukee, WI, Best paper award, June 5-7 2014.

[9] Z. Lin, M. Brouchke, and B. Francis. Local control strategies for groups of mobile autonomous agents. *IEEE Trans. on Automatic Control*, 49:622–629, 2004.

[10] Z. Lin, B. Francis, and M. Maggiore. State agreement for continuous-time coupled nonlinear systems. *SIAM Journal on Control and Optimization*, 46:288–307, 2007.

[11] L. Moreau. Stability of multiagent systems with time-dependent communication links. *IEEE Transactions on Automatic Control*, 50:169–182, 2005.

[12] A. Papachristodoulou, A. Jadbabaie, and U. Munz. Effects of delay in multi-agent consensus and oscillator synchronization. *IEEE Trans. Automa. Control*, 55:1471–1477, 2010.

[13] Z. Qu. *Cooperative Control of Dynamical Systems*. Springer-Verlag, London, 2009.

[14] Z. Qu. Cooperative control of networked nonlinear systems. In *49th IEEE Conf. on Dec. and Ctrl*, pages 3200–3207, Atlanta, GA, Dec 2010.

[15] Z. Qu and M. Simaan. An analytic solution to the optimal design of information structure and cooperative control in networked systems. In *51th IEEE Conf. on Dec. and Ctrl*, pages 4015–4022, Maui, HI, Dec 2012.

[16] Z. Qu, J. Wang, and R. A. Hull. Cooperative control of dynamical systems with application to autonomous vehicles. *IEEE Transactions on Automatic Control*, 53:894–911, 2008.

[17] W. Ren and R. W. Beard. Consensus seeking in multiagent systems under dynamically changing interaction topologies. *IEEE Transactions on Automatic Control*, 50:655–661, 2005.

[18] W. Ren and R. W. Beard. *Distributed Consensus in Multi-vehicle Cooperative Control*. Springer-Verlag, London, 2008.

[19] R. O. Saber, J. Alex Fax, and R. M. Murray. Consensus and cooperation in networked multi-agent systems. *Proceedings of the IEEE*, 95:215–233, 2007.

[20] R. O. Saber and R. M. Murray. Consensus problems in networks of agents with switching topology and time-delays. *IEEE Transactions on Automatic Control*, 49:1520–1533, 2004.

[21] G. N. Saridis and C. G. Lee. An approximation theory of optimal control for trainable manipulators. *IEEE Trans. on Systems, Man and Cybernetics*, 9:152–159, 1979.

[22] E. Semsar-Kazerooni and K. Khorasani. Optimal consensus algorithms for cooperative team of agents subject to partial information. *Automatica*, 44:2766–2777, 2008.

[23] H. G. Tanner, A. Jadbabaie, and G. J. Pappas. Flocking in fixed and switching networks. *IEEE Transactions on Automatic Control*, 52:863–868, 2007.

[24] K. G. Vamvoudakis, F. L. Lewis, and G. R. Hudas. Multi-agent differential graphical games: online adaptive learning solution for synchronization with optimality. *Automatica*, 48:1598–1611, 2012.

[25] D. Vrabie and F. L. Lewis. Neural network approach to continuous-time direct adaptive optimal control for partially-unknown nonlinear systems. *Neural Networks*, 22:237–246, 2009.

[26] J. Wang, M. Obeng, Z. Qu, T. Yang, G. Staskevich, and B. Abbe. Discontinuous cooperative control for consensus of multiagent systems with switching topologies and time-delays. In *the 53th IEEE Conf. on Dec. and Ctrl*, Florence, Italy, Dec 2013.

[27] J. Wang, M. Obeng, T. Yang, G. Staskevich, and B. Abbe. Formation control of multiple nonholonomic mobile robots with limited information of a desired trajectory. In *IEEE International Conference on Electro/Information Technology*, Milwaukee, WI, June 5-7 2014.

[28] J. Wang, Z. Qu, C.M.Ihlefeld, and R. A. Hull. A control-design-based solution to robotic ecology: Autonomy of achieving cooperative behavior from a high-level astronaut command. *Autonomous Robots*, 20:97–112, 2006.

[29] J. Wang, Z. Qu, and M. Obeng. A distributed cooperative steering control with application to nonholonomic robots. In *49th IEEE Conf. on Dec. and Ctrl*, pages 4571–4576, Atlanta, GA, Dec 2010.

[30] J. Wang, C. Smith, G. Staskevich, and B. Abbe. A distributed deployment algorithm for mobile robotic agents with limited sensing/communication ranges. In *IEEE International Conference on Electro/Information Technology*, Naperville, IL, May 21-23 2015.

[31] J. Wang, T. Yang, G. Staskevich, and B. Abbe. Approximate policy iteration for cooperative control of multiagent systems under limited sensing/communications. In *IASTED International Conference on Computational Intelligence*, Innsbruck, Austria, Feb 16-17 2015.

APPENDIX A – Publications and Presentations

The research results have been reported and published in a number of IEEE Conferences. Below is a list of publications, presentations and journal papers under preparation based on the research in this project.

Publications

- Network Controllability Analysis

  1. J. Wang, M. Obeng, Z. Qu, T. Yang, G. Staskevich, and B. Abbe, "Discontinuous cooperative control for consensus of multiagent systems with switching topologies and time-delays," in *53rd IEEE Decision and Control Conference*, Florence, Italy, Dec 2013.

- Optimal Multiagent Coordination

  2. J. Wang, T. Yang, G. Staskevich, and B. Abbe, "Approximate policy iteration for cooperative control of multiagent systems under limited sensing/communications", 2015 IASTED International Conference on Computational Intelligence, Feb 16-17, 2015, Innsbruck, Austria.
  3. J. Wang, T. Yang, G. Staskevich, and B. Abbe, "Approximately adaptive neural cooperative control for multiagent systems with performance guarantee", under revision for journal *IEEE Trans. on Cybernetics*, 2015.
  4. J. Wang, G. Staskevich, and B. Abbe, "Adaptive consensus trajectory tracking for a class of uncertain multiagent systems", *under preparation for journal submission*.
  5. J. Wang, T. Yang, G. Staskevich, and B. Abbe, "Cooperative Q Learning Control of Multiagent Systems", *under preparation for conference submission*.

- Multiagent Coordination Applications: Simulation and Experiments

6. J. Shao, J. Wang, and T. Yang, "Experimental Validation of Distributed Cooperative Control for Mobile Agents with Switching Topologies and Time-Delays", in *IEEE International Conference on Electro/Information Technology*, June 5-7, 2014, Milwaukee, WI, USA. **Best paper award**.
7. J. Wang, M. Obeng, T. Yang, G. Staskevich, and B. Abbe, "Formation Control of Multiple Nonholonomic Mobile Robots with Limited Information of a Desired Trajectory", in *IEEE International Conference on Electro/Information Technology*, June 5-7, 2014, Milwaukee, WI, USA.
8. C. Smith, D. VanDeWater, and J. Wang, "Optimal deployment of mobile robotic agents with limited sensing capabilities", *24th Annual Argonne Symposium for undergraduate/graduate*, Abstract, Argonne National Lab, IL, Nov.7,2014.
9. J. Wang, C. Smith, G. Staskevich, and B. Abbe, "A distributed deployment algorithm for mobile robotic agents with limited sensing/communication ranges", *IEEE International Conference on Electro/Information Technology*, May 21-23, 2015, Naperville, IL, USA.
10. J. Wang, T. Yang, J. Shao, G. Staskevich, and B. Abbe, "Distributed formation control of nonholonomic robots: theory and experiments", *under preparation for journal submission*.

Presentations

- J. Wang, M. Obeng, Z. Qu, T. Yang, G. Staskevich, and B. Abbe, "Discontinuous cooperative control for consensus of multiagent systems with switching topologies and time-delays," in *53rd IEEE Decision and Control Conference*, Florence, Italy, Dec 10- 13, 2013.
- J. Shao, J. Wang, and T. Yang, "Experimental Validation of Distributed Cooperative Control for Mobile Agents with Switching Topologies and Time-Delays", in *IEEE International Conference on Electro/Information Technology*, Milwaukee, WI, USA, June 5-7, 2014. **Best paper award**.
- J. Wang, M. Obeng, T. Yang, G. Staskevich, and B. Abbe, "Formation Control of Multiple Nonholonomic Mobile Robots with Limited Information of a Desired Trajectory", in *IEEE International Conference on Electro/Information Technology*, Milwaukee, WI, USA, June 5-7, 2014.
- C. Smith, D. VanDeWater, and J. Wang, "Optimal deployment of mobile robotic agents with limited sensing capabilities", *24th Annual Argonne Symposium for undergraduate/graduate*, Abstract, Argonne National Lab, IL, Nov.7,2014.
- J. Shao, "Decentralized Cooperative Control for Mobile Agents with Switching Topologies and Time-Delays", *ERAU College of Engineering Industrial Advisory Board meeting*, Daytona Beach, FL, Nov. 3, 2014.

- S. Joseph-Ellison, Q. Zhou, Z. Fu, "System Architecture for a Fleet of Autonomous Underwater Vehicles", *2015 SIAM Conference on Computational Science and Engineering*, one of 12 invited student chapter presentations worldwide, Salt Lake City, Utah, March 15, 2015.
- J. Wang, C. Smith, G. Staskevich, and B. Abbe, "A distributed deployment algorithm for mobile robotic agents with limited sensing/communication ranges", *IEEE International Conference on Electro/Information Technology*, May 23, 2015, Naperville, IL, USA.

APPENDIX B – Copies of Papers

A Distributed Deployment Algorithm for Mobile Robotic Agents with Limited Sensing/Communication Ranges

Jing Wang, Christopher Smith, Gennady Staskevich, and Brian Abbe

Abstract— **In this paper, we propose a distributed deployment algorithm for a group of mobile robots to cover a convex region. The individual mobile robot considered has kinematic constraints, and may only exchange information locally with its neighboring counterparts due to its limited sensing/communication range. The proposed deployment algorithm iteratively updates the Voronoi partition through local information exchange, and then moves toward its centroid based on centroid-drive control algorithms. Particularly, in addition to gradient-based centroid-drive control algorithm in which input-output linearization has been applied to robot model, a new algorithm based on distributed consensus is proposed to directly address the kinematic constraint associated with robot model. Simulation results are provided to illustrate the effectiveness of the proposed algorithm.**

I. Introduction

Coverage control aims to address the issue of deployment of sensor networks for tasks like monitoring an environment, environment modeling, search and rescue, and so on [7][2]. In recent years, mobile autonomous agents have been applied in the construction of mobile sensor networks [3][6][15][8] due to their flexibility and resilience to dynamically changing environments.

Generically, in solving coverage control problem using multiple mobile agents, two fundamental issues have to be addressed: i) how to optimally assign subregions for each agent based on sensing/communication capabilities, online updates about importance of the region to be surveyed, and information exchange among agents; ii) how to design control algorithms to drive agents to the desired deployment locations. In addressing problem i), the typical method is to generate Voronoi diagram for agents by minimizing certain cost functions related to distances between the deployment positions and the measuring points [3]. The importance of region can be embedded into cost function through density function, which could be analytically unknown and can be adaptively estimated using measurement data [15]. For problem ii), the standard gradient-based control can be used if agents assume linear dynamics.

In this paper, we propose a distributed deployment algorithm for a network of mobile robotic agents with kinematic constraints, which is a common characteristic for typical mobile robots such as differential drive robots and car-like robots. In addition, we consider more realistic scenarios of that robots have limited sensing/communication ranges, and solving deployment problem through local information exchange among agents. The proposed work relies on the distributed coordination/consensus results for multiagent systems [11][10][12]. Distributed coordination of multiagent systems has been an active research topics in recent years. Fruitful results are available for the study of cooperative control design and network connectivity analysis [4][12][14][11]. In [11], we proposed sequential completeness condition for sensing/communication matrix sequences connectivity for ensuring coordination of multiple dynamical systems. On the other hand, cooperative control and formation control have been studied extensively for both linear and nonlinear dynamical systems [14][5]. However, there is still a lot of work to be done for dealing with coordination control of mobile robots with kinematic constraints.

The proposed deployment control algorithm in this paper follows a two-step strategy. First, at each time instant, Voronoi partition for each robot is generated based on robot's current position as well as the positions of robots in its communication range. Then control algorithms are designed to drive robots to centroids of Voronoi partitions. In particular, a distributed centroid-drive algorithm is proposed by explicitly taking into account kinematic model constraints for robots. Unicycle robots are used, and the distributed control is based on the transformation of unicycle into chained form [17][16]. Under the assumption of robots maintaining a sequentially complete communication topology, the proposed distributed deployment control algorithm solves the posed coverage control problem. Simulation results are included to illustrate the effectiveness of the proposed design.

II. Problem Formulation

In this paper, we shall consider the problem of deploying a fixed number of mobile robotic agents in a given convex environment Q. An illustration example is shown in figure 1, in which three robots start from some initial positions, and through coordination eventually move to points $[1,3]^T$, $[2,1]^T$, and $[3,3]^T$, respectively, to cover a square area $Q = 4\times 4\ unit^2$. Each robot has the sensing range $r_s = 2.2\ unit$.

To solve the autonomous deployment problem, we make the following assumptions without loss of generality:

- The robots have the knowledge of the area to be covered and sensed.
- The robots have limited sensing ranges r_s, and limited communication ranges r_c. That is, only points in a circle centered at the current robot's position and of radius r_s can be sensed by the robot. In addition, at time t,

J. Wang is with Faculty of Electrical and Computer Engineering, Bradley University, Peoria, IL, 61625. C. Smith is a graduate student with Department of Electrical and Computer Engineering, Bradley University, Peoria, IL, 61625. G. Staskevich and B. Abbe are with AFRL/RISC, Rome, NY, 13441. Email: jingwang@fsmail.bradley.edu This work was supported by the Air Force Research Laboratory FA8750-13-1-0109.

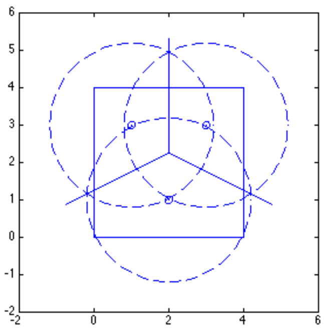

Fig. 1. Deployment of 3 robots in a square area $Q = 4 \times 4\ unit^2$ with sensing range $r_s = 2.2\ unit$

robot i can communicate with its neighboring robot j, $j \in \mathcal{N}_i(t) = \{j | d_{ij} \leq r_c\}$, where d_{ij} is the distance between the ith robot and the jth robot.

- For a given region, there are enough number of n mobile robotic agents to completely cover the area.

To this end, the multiagent coverage control problem is formulated as designing a distributed deployment control algorithm to move the robots towards the centroid of the corresponding partitioned regions based on the minimization of certain coverage cost functions. Under the aforementioned assumptions, the coverage control problem has at least one solution. In this paper, a new paradigm is proposed to deploy the robots by assuming limited sensing and communication ranges.

A. Robot Modeling

The kinematic model of mobile robotic agent carrying sensors is described by the following equations:

$$\begin{aligned} \dot{x}_i &= v_i \cos\theta_i, \\ \dot{y}_i &= v_i \sin\theta_i, \\ \dot{\theta}_i &= \omega_i \end{aligned} \tag{1}$$

where $i \in \Omega \triangleq \{1, \cdots, n\}$, $p_i = [x_i, y_i]^T \in \Re^2$ denotes the ith robot's position, θ_i is the orientation, $v_i \in \Re$ driving velocity, and $\omega_i \in \Re$ the steering velocity. The optimal coverage control problem is then defined as designing distributed cooperative control v_i and ω_i such that agents converge to optimal positions p_i^* by minimizing certain cost function.

Remark 2.1: The model (1) has the so-called nonholonomic constraints [9]. For such a system, there is no continuous state feedback control to solve its stabilization problem due to the violation of Brockett's necessary condition [1]. Therefore, it becomes even more challenging to address the optimal deployment problem of multiple mobile agents with nonholonomic constrains.

III. Proposed Deployment Algorithm

In this section we present a distributed deployment algorithm for mobile robotic agents with limited sensing/communication ranges. The proposed deployment algorithm is a recursive one. At each sampling time instant, each robot first computes its Voronoi cell based on its communication with neighboring robots, then determine the centroid of its Voronoi region, and then moves towards it by employing a distributed coordination algorithm.

A. Voronoi Partition

In solving coverage control problem for sensor networks, Voronoi diagram has been popular in generating the deployment positions for sensor nodes [3]. In what follows, we describe the basic idea of Voronoi partition based coverage control for mobile robots.

Let us denote an arbitrary point in the region Q as q. At each sampling time instant, the agents will be able to generate the Voronoi partition of Q. That is, for agent i at position p_i, its Voronoi region satisfies

$$V_i = \{q \in Q | \|q - p_i\| \leq \|q - p_j\|, \forall j \neq i\} \tag{2}$$

Define cost function over the region as

$$J(p_1, \cdots, p_n) = \sum_i^n \int_{V_i} \frac{1}{2} \|q - p_i\|^2 \phi(q) dq \tag{3}$$

where $\phi(q)$ is a weighting function of importance over Q. The distance function $\frac{1}{2}\|q - p_i\|^2$ is included in the cost function for the consideration of reducing energy consumed by a sensor's transceiver because it is generally a function of distance. In addition, the reliability of the information at q measured by robot at p_i will degrade with the increase of distance $\|q - p_i\|^2$.

At each sampling time instant, the generation of Voronoi region V_i for robot i is based on the robots in its neighboring set $\mathcal{N}_i$. That is, robot i can only use the position information of the robots in its communication range r_c to compute V_i. This is a realistic situation since during the motion, the robot could move in or out the communication range which is limited. It is apparent that, by only considering the robots in its communication range, the obtained Voronoi partition could be different. For instance, consider a robot at location $[2, 1]^T$ computing its Voronoi region for a square area $Q = 4 \times 4\ unit^2$. Figure 2 and figure 3 show the resulting Voronoi region under 3 neighboring robots and 2 neighboring robots, respectively.

Once the Voronoi region is obtained, a simple control to drive the robot to the centroid of the Voronoi region is to follow negative gradient of cost function J, that is,

$$-\frac{\partial J}{\partial p_i} = -\int_{V_i} (q - p_i)\phi(q) dq$$

However, as discussed before, the kinematic model in (1) is nonlinear and may not be able to follow the negative gradient due to velocity constraints. A simple way to avoid this issue is to conduct input/output linearization by choosing a reference point off the robot center (x_i, y_i), that is, let the cartesian coordinates of the off-center reference point be

$$p_{i1} = x_i + b\cos\theta_i \tag{4}$$

$$p_{i2} = y_i + b\sin\theta_i \tag{5}$$

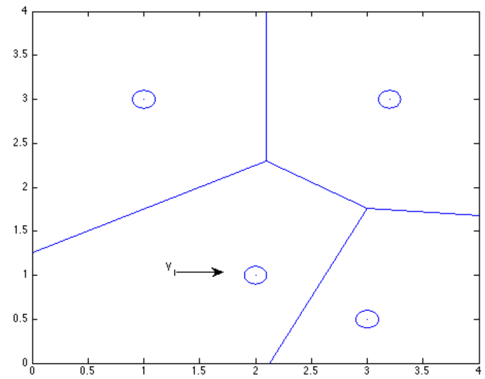

Fig. 2. Voronoi region for robot at $[2,1]^T$ with 3 neighboring robots

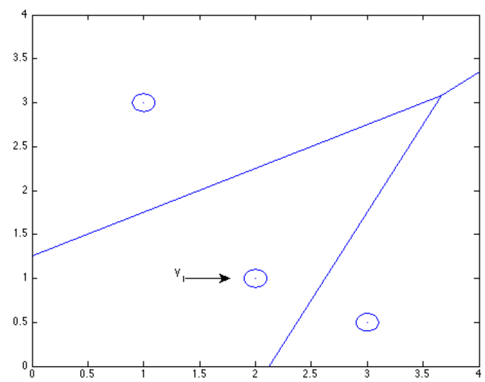

Fig. 3. Voronoi region for robot at $[2,1]^T$ with 2 neighboring robots

where $b > 0$ is a constant. Differentiating (4) and (5) with respect to time, we have

$$\begin{aligned} \begin{bmatrix} \dot{p}_{i1} \\ \dot{p}_{i2} \end{bmatrix} &= \begin{bmatrix} \cos\theta & -b\sin\theta \\ \sin\theta & b\cos\theta \end{bmatrix} \begin{bmatrix} v_i \\ \omega_i \end{bmatrix} \\ &\triangleq T(\theta) \begin{bmatrix} v_i \\ \omega_i \end{bmatrix} \triangleq \begin{bmatrix} u_{i1} \\ u_{i2} \end{bmatrix} \end{aligned} \tag{6}$$

To this end, the distributed deployment control for robot i is given by

$$\begin{aligned} u_i &\triangleq \begin{bmatrix} u_{i1} \\ u_{i2} \end{bmatrix} = -\frac{\partial J}{\partial p_i} \\ &= -\int_{V_i} (q - p_i)\phi(q)dq = -M_{V_i}(C_{V_i} - p_i) \end{aligned} \tag{7}$$

where mass M_{V_i} is given by

$$M_{V_i} = \int_{V_i} \phi(q)dq \tag{8}$$

the first moment

$$L_{V_i} = \int_{V_i} q\phi(q)dq \tag{9}$$

and the centroid

$$C_{V_i} = \frac{L_{V_i}}{M_{V_i}} \tag{10}$$

Once u_i is obtained, the control inputs v_i and ω_i can be calculated by using inverse input transformation given below:

$$\begin{bmatrix} v_i \\ \omega_i \end{bmatrix} = T^{-1}(\theta) \begin{bmatrix} u_{i1} \\ u_{i2} \end{bmatrix}.$$

B. Deployment Algorithm Based on Distributed Consensus

In this subsection, a distributed deployment algorithm by directly dealing with nonlinear model in (1) is presented. That is, we propose a new control to drive robots to Voronoi centroids based on distributed consensus algorithms.

To start, we first convert (1) into the following canonical chained form

$$\begin{aligned} \dot{z}_{i1} &= u_{i1} \\ \dot{z}_{i2} &= u_{i2} \\ \dot{z}_{i3} &= z_{i2}u_{i1} \end{aligned} \tag{11}$$

by using the state and input transformations defined below

$$z_{i1} = x_i, \quad z_{i2} = \tan\theta_i, \quad z_{i3} = y_i, \tag{12}$$

$$u_{i1} = v_i \cos\theta_i, \quad u_{i2} = \frac{\omega_i}{\cos^2\theta_i}. \tag{13}$$

The controls u_{i1} and u_{i2} will be designed and the corresponding v_i and ω_i can be obtained through the inverse transformation of (13).

To facilitate the design, we apply a binary matrix $C(t)$ to describe the time-varying communication topologies among robots, that is, given a time sequence $\{t^s_\eta : \eta = 0, 1, \cdots\}$, $C(t)$ is defined by

$$C(t) = \begin{bmatrix} c_{11} & c_{12}(t) & \cdots & c_{1n}(t) \\ c_{21}(t) & c_{22} & \cdots & c_{2n}(t) \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1}(t) & c_{n2}(t) & \cdots & c_{nn} \end{bmatrix}, \tag{14}$$

with $C(t) = C(t^s_\eta), \forall t \in [t^s_\eta, t^s_{\eta+1})$, where $c_{ii} \equiv 1$; $c_{ij}(t) = 1$ if the jth robot is in the sensing/communication range of the ith robot at time t, and $c_{ij} = 0$ if otherwise; and $t^s_0 \triangleq t_0$. It can be assumed without loss of any generality that $0 < \underline{c}_t \leq t^s_{\eta+1} - t^s_\eta \leq \overline{c}_t < \infty$, where $\underline{c}_t$ and $\overline{c}_t$ are constant bounds.

Remark 3.1: At each sampling time instant, finite-time steering control can be used to move robots to the centroids of Voronoi regions. In what follows, in order to further improve the robustness against measurement errors, we propose a distributed coordination algorithm based on information exchange among neighboring robots.

The proposed design is based on distributed consensus idea, which requires that the communication topologies defined by (14) satisfy *sequential completeness* condition [11][17]. The *sequential completeness* condition describe the least required condition on network connectivity for cooperative control design, which is equivalent to the existence of a spanning tree introduced in [12].

Assumption 3.1: The group of robots defined in (1) has a sequentially complete communication network.

In what follows, we present the distributed deployment algorithm under assumption 3.1. Define an infinite sequence of time instants $\{t_0 + kT_s\}$ for $k \in \mathcal{N} \overset{\triangle}{=} \{0, 1, \cdots\}$ and with sampling time $0 < T_s \leq \underline{c}_t$. The control inputs will be updated according to the sampling time instants. For notational convenience, $z(t_0 + kT_s)$ is simplified as $z(k)$ for any variable z. At each time instant k, the Voronoi partition V_i is obtained for robot i based on local information, then its centroid $[C_{v_i,x}(k), C_{v_i,y}(k)]^T$ is computed, and then the distributed centroid control algorithm is used for moving robots to the desired positions. The proposed distributed centroid control algorithm is summarized as follows.

Let the distributed centroid control be for $t \in [t_0 + kT_s, t_0 + (k+1)T_s)$

$$u_{i1}(t) = a_{i1}^k + a_{i2}^k \sin\omega(t - t_0 - kT_s) \tag{15}$$

$$u_{i2}(t) = b_{i1}^k + b_{i2}^k \cos\omega(t - t_0 - kT_s) \tag{16}$$

where $\omega = \frac{2\pi}{T_s}$, $a_{i2}^k \neq 0$ can be any constant, and

$$a_{i1}^k = \frac{1}{T_s}\sum_{j=1}^{n} G_{ij}(k)[x_j(k) - x_i(k) - C_{v_j,x}(k) + C_{v_i,x}(k)], \tag{17}$$

$$b_{i1}^k = \frac{1}{T_s}\sum_{j=1}^{n} G_{ij}(k)[z_{j2}(k) - z_{i2}(k)], \tag{18}$$

$$b_{i2}^k = \frac{2\omega}{a_{i2}^k T_s}\left[\sum_{j=1}^{n} G_{ij}(k)[y_j(k) - y_i(k) - C_{v_j,y}(k) + C_{v_i,y}(k)] - \frac{a_{i1}^k b_{i1}^k T_s^2}{2} - a_{i1}^k z_{i2}(k) T_s + \frac{a_{i2}^k b_{i1}^k T_s}{\omega}\right]. \tag{19}$$

with

$$G_{ij}(k) = \frac{c_{ij}(k)}{\sum_{\eta=1}^{n} c_{i\eta}(k)}, \quad j = 1, \cdots, n. \tag{20}$$

In the use of algorithms (15) and (16), the centroid $[C_{v_i,x}(k), C_{v_i,y}(k)]^T$ at each step will be generated using (10).

In summary, the proposed distributed deployment algorithm is given in Algorithm 1.

Algorithm 1 Distributed Deployment Algorithm

1: Let $k = 0$. Given initial states $p_i(k)$, $c_{ij}p_j(k)$, calculate the initial Voronoi partition $V_i(k)$.
2: Compute the centroid $[C_{v_i,x}(k), C_{v_i,y}(k)]^T$ using (10).
3: Employ control (7) or (15)-(16).
4: Let $k \leftarrow k + 1$, and go to step 2, until

$$(C_{v_i,x}(k+1) - C_{v_i,x}(k))^2 + (C_{v_i,y}(k+1) - C_{v_i,y}(k))^2) \leq \epsilon,$$

where $\epsilon > 0$ is a sufficiently small predefined threshold.

IV. Simulation

In this section, we simulate the proposed distributed deployment algorithm. Consider first the case with 5 mobile robotic agents, and we assume fully connected communication topology. That is, at each time instant, each robot has the position information of the rest members in the group. Figure 4 and 5 illustrate the initial location with Voronoi partition and the final position with Voronoi partition, respectively. Figure 6 illustrates of the evolution of the robots.

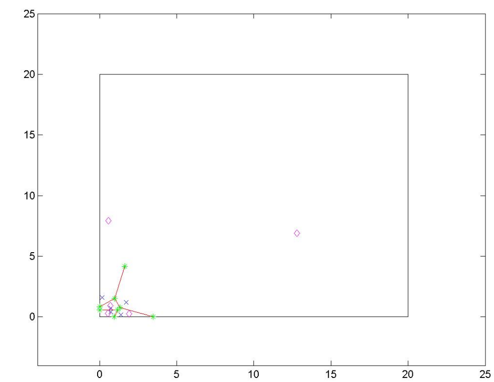

Fig. 4. Initial location and Voronoi partition

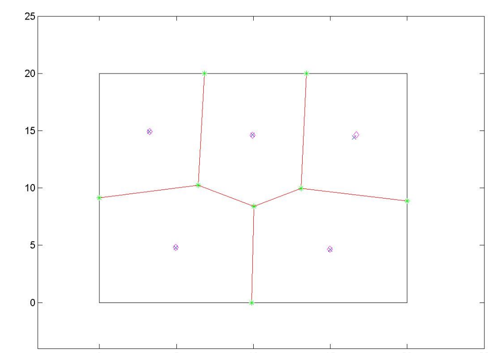

Fig. 5. Final location and Voronoi partition

In the 2nd case, we consider 10 robots with limited communication ranges. Assume that the initial communication topology is defined by

$$C(0) = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

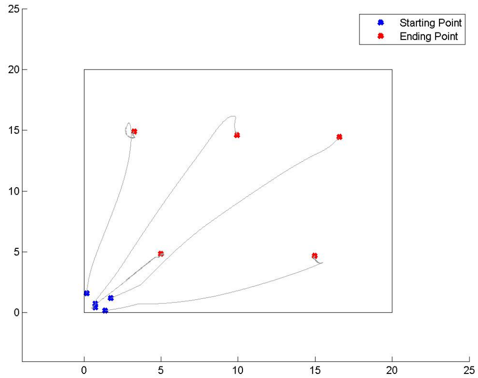

Fig. 6. Evolution of the robots

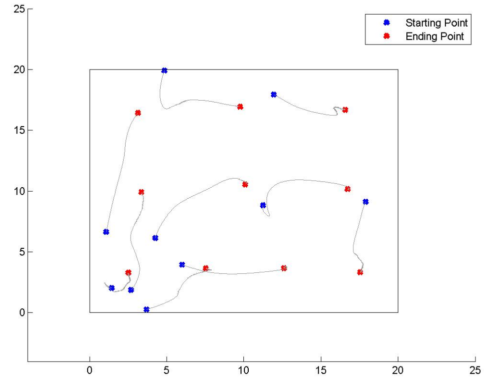

Fig. 9. Evolution of the robots

and changes subsequently based on system evolution. Figure 7 and 8 illustrate the initial location with Voronoi partition and the final position with Voronoi partition, respectively. Figure 9 illustrates of the evolution of the robots.

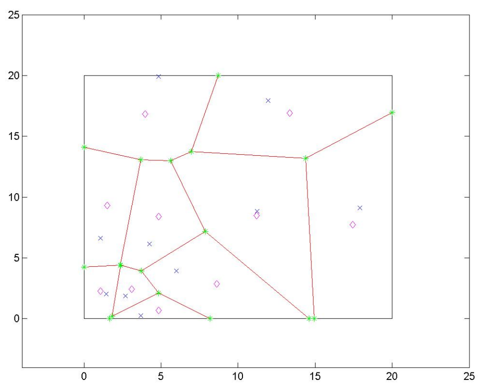

Fig. 7. Initial location and Voronoi partition

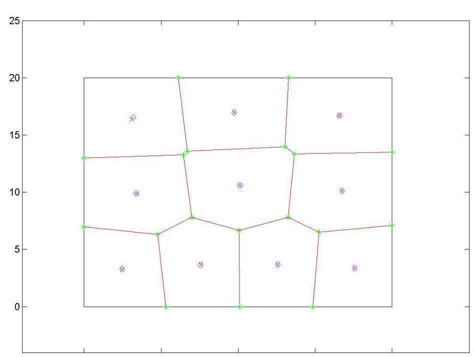

Fig. 8. Final location and Voronoi partition

V. CONCLUSION

In this paper, we proposed a distributed deployment algorithm for solving the coverage control problem of mobile robotic agents with inherent kinematic constraints. The proposed design assumes the limited sensing/communication capabilities, and the generation of Voronoi partition for each robot as well as the centroid-drive control are based on local information exchange among agents. Simulation results validated the effectiveness of the proposed design. Future work will be focused on experimental validation of the proposed algorithm by using Kilobots [13].

REFERENCES

[1] R. W. Brockett, "Asymptotic stability and feedback stabilization," in *Differential Geometric Control Theory*, vol. R. W. Brockett, R. S. Millman, and H. J. Sussmann, Eds, 1983, pp. 181–191.

[2] H. Choset, "Coverage for robotics - a survey of recent results," in *Annals of Mathematics and Artificial Intelligence*, vol. 31, 2001, pp. 113–126.

[3] J. Cortes, S. Martinez, T. Karatas, and F. Bullo, "Coverage control for mobile sensing networks," *IEEE Trans. on Robotics and Automation*, vol. 20, pp. 243–255, 2004.

[4] A. Jadbabaie, J. Lin, and A. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Trans. on Automatic Control*, vol. 48, pp. 988–1001, 2003.

[5] Z. Lin, B. Francis, and M. Maggiore, "State agreement for continuous-time coupled nonlinear systems," *SIAM Journal on Control and Optimization*, vol. 46, pp. 288–307, 2007.

[6] Z. Lin, S. Zhang, and G. Yan, "An incremental deployment algorithm for wireless sensor networks using one or multiple autonomous agents," *Ad Hoc Networks*, vol. 11, pp. 355–367, 2013.

[7] K. Lynch, I. Schwartz, P. Yang, and R. A. Freeman, "Decentralized environmental modeling by mobile sensor networks," *IEEE Trans. on Robotics*, vol. 24, pp. 710–724, 2008.

[8] S. Martinez and F. Bullo, "Optimal sensor placement and motion coordination for target tracking," *Automatica*, vol. 42, pp. 661–668, 2006.

[9] R. M. Murray and S. S. Sastry, "Nonholonomic motion planning: Steering using sinusoids," *IEEE Trans. on Auto. Contr.*, vol. 38, pp. 700–716, 1993.

[10] Z. Qu, *Cooperative Control of Dynamical Systems*. London: Springer-Verlag, 2009.

[11] Z. Qu, J. Wang, and R. A. Hull, "Cooperative control of dynamical systems with application to autonomous vehicles," *IEEE Transactions on Automatic Control*, vol. 53, pp. 894–911, 2008.

[12] W. Ren and R. W. Beard, *Distributed Consensus in Multi-vehicle Cooperative Control*. London: Springer-Verlag, 2008.

[13] M. Rubenstein, A. Cornejo, and R. Nagpal, "Programmable self-assembly in a thousand-robot swarm," *Science*, vol. 345, pp. 795–799, 2014.

[14] R. O. Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, pp. 215–233, 2007.

[15] M. Schwager, D. Rus, and J.-J. Slotine, "Decentralized, adaptive coverage control for networked robots," *The International Journal of Robotics Research*, vol. 28, pp. 357–375, 2009.

[16] J. Wang, M. Obeng, T. Yang, G. Staskevich, and B. Abbe, "Formation control of multiple nonholonomic mobile robots with limited information of a desired trajectory," in *IEEE International Conference on Electro/Information Technology*, Milwaukee, WI, June 2014.

[17] J. Wang, Z. Qu, and M. Obeng, "A distributed cooperative steering control with application to nonholonomic robots," in *49th IEEE Conf. on Dec. and Ctrl*, Atlanta, GA, Dec 2010, pp. 4571–4576.

52nd IEEE Conference on Decision and Control
December 10-13, 2013. Florence, Italy

Discontinuous Cooperative Control for Consensus of Multiagent Systems with Switching Topologies and Time-Delays

Jing Wang, Morrison Obeng, Zhihua Qu, Tianyu Yang, Gennady Staskevich, and Brian Abbe

Abstract— **In this paper, we propose a discontinuous cooperative control for consensus of multiagent systems with directed and switching sensing/communication topologies and time-delays. By introducing a new design for nonlinear cooperative control gains, multiagent system consensus can be guaranteed in the presence of switching topologies and time-delays. System convergence analysis is done by employing a new contraction mapping method. Simulation examples are provided to illustrate the effectiveness of the proposed design.**

I. Introduction

Cooperative control of multiagent systems has attracted a great deal of attention in recent years [18][14][1][19]. Multiagent systems are generically defined as a group of dynamical systems in which certain emergent behaviors are exhibited through the local interaction of group members that individually have the capability of self-operating. Fundamentally, the key issues in engineered multiagent systems are the study of network controllability and the design of distributed cooperative control. In terms of network controllability, the objective is to figure out the connectivity conditions on sensor/communication topologies of the network for achieving consensus behavior. In [8][20], the condition is obtained for composite undirected graphs which need to be connected. Extensions were made in [17][9] to the case with directed graphs, and the less restrictive conditions are stated as that there exists a spanning tree or the network is strongly connected periodically. Complement to the aforementioned graph-theoretical methods, a matrix-theoretical framework is developed in [16] to deal with the high-order systems with arbitrary but finite relative degrees. It is shown that network controllability is ensured if and only if the sensing/communication network is sequentially complete.

The design of cooperative control is closely related to system dynamics. For linear systems, the results in [8][17][9][6][20] are developed for the first-order integrator model, in [22] for double integrator model, and in [16][23] for high-order linear model. For nonlinear systems, the problem becomes complicated since network controllability may not render the direct design of cooperative control and system dynamics have to be explicitly taken into account. In [11], a solution is obtained by convexity analysis for a class of discrete-time nonlinear systems. The continuous-time nonlinear systems are also addressed such as in [10][13][15]. Particularly, a subtangentiality condition on the vector fields is identified in [10]. In [13], the local passivity condition is imposed on nonlinear functions in the system dynamics, while a diagonally quasi-linear functions of positive gains is introduced in [15]. In addition, time delays are literally analyzed in [13][15] for continuous-time nonlinear systems. It should be noted that the results in [10][13][15][24] are for nonlinear systems with smooth dynamics. There also appeared some pioneering work on consensus of systems with discontinuous dynamics [5][3][2][7][4] by using the tools from nonsmooth analysis [21][12]. Discontinuous control law was proposed for the coordination of nonholonomic mobile robots in [5]. The finite-time semistable concept was introduced in [7] for a class of switched rendezvous protocols. The results in [3][2] addressed the distributed estimation and tracking problem using a variable structure approach, and a binary consensus control protocol was designed in [4] via a pin node.

In this paper, we propose a new discontinuous cooperative control design for multiagent systems with switching and directed sensing/communication topologies. The case in the presence of sensing/communication delays is also rigorously addressed. Particularly, we developed a contraction mapping method for the consensus analysis of multiagent systems under the proposed discontinuous cooperative control. The proposed discontinuous cooperative control design provides a possible way to address the cooperative control problem with more complicated system dynamics, and enriches the disposal for cooperative control protocols. The contributions of the paper are two-fold. First, it reveals that network controllability condition does not guarantee the consensus in the presence of discontinuous system dynamics. Second, it is rigorously proved that through designing nonlinear piecewise control gains, the convergence can be ensured for multiagent systems with switching topologies and time-delays under the least-restrictive network controllability condition of that the system sensing/communication topologies are sequentially complete. Simulation examples are provided to illustrate the effectiveness of the proposed design.

II. Problem Formulation

Consider a multiagent system which has n members and each agent assumes the single-integrator dynamics

$$\dot{x}_i = u_i, \tag{1}$$

J. Wang and M. Obeng are with Faculty of Computer Engineering, School of Science, Engineering and Mathematics, Bethune-Cookman University, Daytona Beach, FL 32114, USA. Z. Qu is with School of Electrical Engineering and Computer Science, University of Central Florida, Orlando, FL 32816, USA. G. Staskevich and B. Abbe are with AFRL/RISC, Rome, NY, 13441. T. Yang is with Electrical Engineering, Embry-Riddel Aeronautical University, Daytona Beach, FL 32114. Emails: wangj@cookman.edu; qu@eecs.ucf.edu This work was supported by the Air Force Research Laboratory FA8750-13-1-0109.

978-1-4673-5717-3/13/$31.00 ©2013 IEEE

where $i \in \Omega \triangleq \{1, \cdots, n\}$, $x_i(t) \in \Re$ is the state, $u_i \in \Re$ is the control input to be designed. The objective of this paper is to design a discontinuous cooperative control $u_i(t)$ to achieve the consensus of the multiagent system (1), that is,

$$\lim_{t\to\infty} x_i(t) = x^*, \quad \forall i, \tag{2}$$

where x^* is some constant denoting the consensus value.

To reach the consensus, the control design will be based on the sensing/communication information exchange among agents, which can be described by the following sensing/communication matrix and its corresponding time sequence $\{t_k^s : k = 0, 1, \cdots\}$. That is, within time interval $[t_k^s, t_{k+1}^s)$, the sensing/communication topology is assumed to be unchanged.

$$\begin{aligned} S(t_k^s) &= \begin{bmatrix} s_{11} & s_{12}(t_k^s) & \cdots & s_{1q}(t_k^s) \\ s_{21}(t_k^s) & s_{22} & \cdots & s_{2q}(t_k^s) \\ \vdots & \vdots & \ddots & \vdots \\ s_{q1}(t_k^s) & s_{q2}(t_k^s) & \cdots & s_{qq} \end{bmatrix}, \\ S(t) &= S(t_k^s), \quad \forall t \in [t_k^s, t_{k+1}^s), \end{aligned} \tag{3}$$

where $s_{ii} \equiv 1$; $s_{ij}(t) = 1$ if the ith agent can receive the information from the jth agent at time t, and $s_{ij} = 0$ if otherwise; and $t_0^s \triangleq t_0$. The neighbor set of agent i is defined as $\mathcal{N}_i = \{j \in \Omega | s_{ij} \neq 0\}$. We further assume without loss of any generality that $0 < \underline{c}_t \leq t_{k+1}^s - t_k^s \leq \overline{c}_t < \infty$, where $\underline{c}_t$ and $\overline{c}_t$ are constant bounds.

The proposed cooperative control is of the form

$$\begin{aligned} u_i(t) &= \sum_{l=1}^{n} \alpha_{il}(s_{il}(t_k^s), x_l(t_k^s))\text{sgn}(x_l(t) - x_i(t)), \\ & \quad t \in [t_k^s, t_{k+1}^s), \end{aligned} \tag{4}$$

where $\alpha_{il}(\cdot, \cdot)$ is a nonlinear gain to be designed based on the sensing/communication topology $S(t_k^s)$ as well as the available boundary values $x_l(t_k^s)$ if $s_{il}(t_k^s) \neq 0$, and sgn$(\cdot)$ function is defined as

$$\text{sgn}(z) = \begin{cases} 1, & z > 0 \\ 0, & z = 0 \\ -1, & z < 0 \end{cases}$$

III. Main Results

Assume that the multiagent system (1) is operating under switching and directed sensing/communication topologies. That is, sensing/communication matrix $S(t_k^s)$ is changing, and not necessary be symmetric (in general $s_{ij}(t_k^s) \neq s_{ji}(t_k^s)$).

To proceed with the design and stability analysis for the closed-loop system under control (4), we introduce the following definitions which are adapted from [16] and describe the standing conditions on sensing and communication.

Definition 3.1: Sensing/communication matrix sequence $\{S(t)\}$ is said to be *sequentially lower-triangularly complete* if it is sequentially lower-triangular and in every row i of its lower triangular canonical form, there is at least one $j < i$ such that the corresponding block is uniformly non-vanishing.

Definition 3.2: Sensing/communication matrix sequence $\{S(t)\}$ is said to be *sequentially complete* if the sequence contains an infinite subsequence that is sequentially lower-triangularly complete.

Remark 3.1: The *sequential completeness* concept of the sensing/communication matrix sequence $\{S(t)\}$ was first introduced in [16]. It spells out the least restrictive connectivity condition for sensor/communication network in order to achieve the asymptotically cooperative stability of the overall system. It is equivalent to condition of the existence of a spanning tree in the graph theory [14]. As an example, consider the following communication sequence,

$$S(t_{3k}) = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, S(t_{3k+1}) = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$
$$S(t_{3k+2}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \tag{5}$$

where $k = 0, 1, \cdots$. It is readily verified that the matrix sequence $\{S(t_{3k}),\ S(t_{3k+1}),\ S(t_{3k+2})\}$ is sequentially complete since the sub-sequence consisting of $S(t_{3k})$ and $S(t_{3k+2})$ is sequentially lower triangular complete. $\diamond$

A. Motivating Example

For the cooperative control of multiagent systems (1), if the standard design of $u_i(t)$ is adopted as given below

$$u_i(t) = \sum_{l=1}^{n} \alpha_{il}(s_{il}(t_k^s))(x_l(t) - x_i(t)), t \in [t_k^s, t_{k+1}^s), \tag{6}$$

where

$$\alpha_{il}(t_k^s) = \frac{s_{il}(t_k^s)}{\sum_{j=1} s_{ij}(t_k^s)}, \tag{7}$$

then it has been proved in [16] that the sequential completeness of sensing/communication matrix sequence $\{S(t)\}$ is the necessary and sufficient condition for consensus of multiagent systems. However, under the discontinuous cooperative control (4) proposed in this paper, the sequential completeness of sensing/communication network may no longer ensure the consensus if the gains α_{il} are simply designed using (7). This is illustrated through the following example.

Example 1: Suppose we have 3 agents. Define index set $\Omega = \{1, 2, 3\}$, $\Omega_{\max} = \{i \in \Omega : x_i(t) = x_{\max}(t) \triangleq \max_j x_j(t)\}$, and $\Omega_{\min} = \{i \in \Omega : x_i(t) = x_{\min}(t) \triangleq \min_j x_j(t)\}$.

Assume that at time instant t_0, we have $\Omega_{\min}(t_0) = \{1\}$, and $\Omega_{\max}(t_0) = \{2, 3\}$, and the sensing/communication topologies among three agents switch according to sensing/communication matrices $S(t_{3k}), S(t_{3k+1})$ and $S(t_{3k+2})$ defined in (5).

It can be readily verified that the matrix sequence $S(t_{3k}), S(t_{3k+1}), S(t_{3k+2})$ is sequentially complete. However, the consensus is not guaranteed if the standard gain design for α_{ij} in (7) is applied under control (4). One possible scenario is that according to the sensing/communication

matrix $S(t_0)$, agent 2 receives information from agent 1 and may converge to agent 1 in finite time interval $t_1 - t_0$, thus at time instant t_1, we could have $\Omega_{\min}(t_1) = \{1, 2\}$ and $\Omega_{\max}(t_1) = \{3\}$; similarly, according to $S(t_1)$, agent 1 receives information from agent 3 and may converge to agent 3 in finite time $t_2 - t_1$, thus we may have $\Omega_{\min}(t_2) = \{2\}$ and $\Omega_{\max}(t_2) = \{1, 3\}$; by $S(t_2)$, agent 3 receives information from agent 2 and may converge to agent 2 in finite time $t_3 - t_2$, and we may have $\Omega_{\min}(t_3) = \{2, 3\}$ and $\Omega_{\max}(t_3) = \{1\}$. This pattern will repeat following the periodical sensing/communication matrix sequence $\{S(t_i)\}$. In other words, though within time interval $[t_0, t_3)$, the communication topology is complete, contraction mapping is not established since we have $x_{\max}(t_3) = x_{\max}(t_0)$ and $x_{\min}(t_3) = x_{\min}(t_0)$ from the above analysis. This is further illustrated in figure 1, in which we consider three agents with controls (4) and gain $\alpha_{ij}(t)$ are chosen based on (7), simulation parameters are given as $t_{3k+i} - t_{3k+i-1} = 0.1, i = 1, 2, k = 0, 1, \cdots$, and initial conditions $x_1(t_0) = 0, x_2(t_0) = 0$, and $x_3(t_0) = 0.1$. Apparently, no consensus is reached. $\diamond$

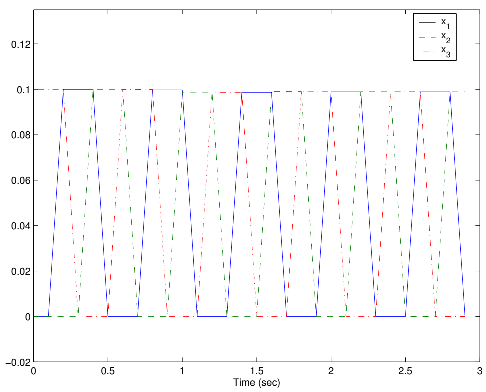

Fig. 1. System responses

In the presence of sensing/communication delays, the cooperative control in (4) becomes

$$u_i(t) = \sum_{l=1}^{n} \alpha_{il}(s_{il}, x_l(t_k^s - \tau_{il}))\operatorname{sgn}(x_l(t - \tau_{il}) - x_i(t)), \quad t \in [t_k^s, t_{k+1}^s), \tag{8}$$

where $\tau_{il} \in [0, r]$ are time delays incurred during transmission with r being the upper bound on latencies of information transmission over the network. In general, multiagent systems with time-delays become more involved. By imposing more stringent network connectivity conditions, such as bi-directional (undirected) sensing/communication, the consensus may still be ensured. However, given the discontinuous cooperative control (8), if control gains α_{ij} are simply chosen according to (7), consensus cannot be guaranteed even under fixed and undirected communication topology as illustrated by the following example. Nonlinear piecewise constant gain $\alpha_{ij}(\cdot)$ needs to be designed to solve the problem.

B. Design and Stability Analysis with Directed and Switching Topologies

As shown in example 1 , standard network topology based control gain design for (4) no longer implies the consensus of multiagent systems, even with the most-restrictive network connectivity condition (that is, fixed and undirected communication). In this subsection, in order to ensure the multiagent systems consensus with control (4) under the least restrictive sensing/communication condition (that is, sequential completeness of $\{S(t_k^s)\}$, we propose a new nonlinear piecewise gain design. The convergence of the overall closed-loop systems is proved by developing a contraction mapping method for multiagent systems.

Theorem 1: Consider the multiagent system (1) under cooperative control (4). Assume that sensing/communication matrix sequence $\{S(t_k^s)\}$ is uniformly sequentially complete*. Let the nonlinear gain α_{lj} be designed as follows: for any agent l,

1) if $x_l(t_k^s) = \max_{j\in\mathcal{N}_l} x_j(t_k^s) = \min_{j\in\mathcal{N}_l} x_j(t_k^s)$, then $\alpha_{lj}(t_k^s)$ can be any bounded positive value.
2) if $x_l(t_k^s) \geq \max_{j\in\mathcal{N}_l} x_j(t_k^s)$, let $\alpha_{lj}(t_k^s)$ be selected to satisfy the inequality
$$0 \leq \sum_{j\in\mathcal{N}_l} \alpha_{lj}(t_k^s) < \frac{x_l(t_k^s) - \min_{j\in\mathcal{N}_l} x_j(t_k^s)}{\bar{c}_t} \tag{9}$$
3) if $x_l(t_k^s) \leq \min_{j\in\mathcal{N}_l} x_j(t_k^s)$, let $\alpha_{lj}(t_k^s)$ be selected to satisfy the inequality
$$0 \leq \sum_{j\in\mathcal{N}_l} \alpha_{lj}(t_k^s) < \frac{\max_{j\in\mathcal{N}_l} x_j(t_k^s) - x_l(t_k^s)}{\bar{c}_t} \tag{10}$$
4) if $\min_{j\in\mathcal{N}_l} x_j(t_k^s) < x_l(t_k^s) < \max_{j\in\mathcal{N}_l} x_j(t_k^s)$, let $\alpha_{lj}(t_k^s)$ be selected to satisfy the inequality
$$0 \leq \sum_{j\in\mathcal{N}_l} \alpha_{lj}(t_k^s) < \min\left(\begin{array}{c} \frac{\max_{j\in\mathcal{N}_l} x_j(t_k^s) - x_l(t_k^s)}{\bar{c}_t}, \\ \frac{x_l(t_k^s) - \min_{j\in\mathcal{N}_l} x_j(t_k^s)}{\bar{c}_t} \end{array} \right) \tag{11}$$

Then consensus of system (1) is asymptotically achieved in the sense of (2).

Proof: See Appendix.

The nonlinear gain design conditions (9) to (11) play a paramount important role for the consensus of multiagent systems (1). Those conditions are easy to be satisfied since for agent l, it only requires the available neighboring state information of agent l in the design of $\alpha_{lj}(t_k^s)$. For instance, to satisfy (9), one simple choice could be

$$\alpha_{lj}(t_k^s) = \frac{x_l(t_k^s) - \min_{j\in\mathcal{N}_l} x_j(t_k^s)}{(|\mathcal{N}_l| + 1)\bar{c}_t}, \quad \forall l \in \mathcal{N}_l \tag{12}$$

where $|\mathcal{N}_l|$ denotes the cardinality of the set $\mathcal{N}_l$. Same selection can be made for satisfying the conditions (10) and (11).

*The time-varying sensing/communication topology is considered here. If the topology becomes fixed after certain time, we can treat it as a special case of switching sensing/communication sequence $S(t_k^s)$ with $\bar{c}_t$ being any positive constants.

C. Multiagent Systems with Time-Delays

The following theorem presents the control design and consensus analysis for multiagent systems with directed switching communications and time-delays.

Theorem 2: Consider the multiagent system (1) under cooperative control (8). Assume that sensing/communication matrix sequence of $\{S(t_k^s)\}$ is sequentially complete. Let the nonlinear gain α_{lj} be designed as follows: for any agent l,

1) if $x_l(t_k^s) = \max_{j\in\mathcal{N}_l} x_j(t_k^s - \tau_{lj}) = \min_{j\in\mathcal{N}_l} x_j(t_k^s - \tau_{lj})$, then $\alpha_{lj}(t_k^s)$ can be any bounded positive values.
2) if $x_l(t_k^s) \geq \max_{j\in\mathcal{N}_l} x_j(t_k^s - \tau_{lj})$, let $\alpha_{lj}(t_k^s)$ be selected to satisfy the inequality

$$0 \leq \sum_{j\in\mathcal{N}_l} \alpha_{lj}(t_k^s) < \frac{x_l(t_k^s) - \min_{j\in\mathcal{N}_l} x_j(t_k^s - \tau_{lj})}{\max\{\bar{c}_t, r\}} \tag{13}$$

3) if $x_l(t_k^s) \leq \min_{j\in\mathcal{N}_l} x_j(t_k^s - \tau_{lj})$, let $\alpha_{lj}(t_k^s)$ be selected to satisfy the inequality

$$0 \leq \sum_{j\in\mathcal{N}_l} \alpha_{lj}(t_k^s) < \frac{\max_{j\in\mathcal{N}_l} x_j(t_k^s - \tau_{lj}) - x_l(t_k^s)}{\max\{\bar{c}_t, r\}} \tag{14}$$

4) if $\min_{j\in\mathcal{N}_l} x_j(t_k^s - \tau_{lj}) < x_l(t_k^s) < \max_{j\in\mathcal{N}_l} x_j(t_k^s - \tau_{lj})$, let $\alpha_{lj}(t_k^s)$ be selected to satisfy the inequality

$$0 \leq \sum_{j\in\mathcal{N}_l} \alpha_{lj}(t_k^s) < \min\left(\begin{array}{c} \frac{\max_{j\in\mathcal{N}_l} x_j(t_k^s - \tau_{lj}) - x_l(t_k^s)}{\max\{\bar{c}_t, r\}}, \\ \frac{x_l(t_k^s) - \min_{j\in\mathcal{N}_l} x_j(t_k^s - \tau_{lj})}{\max\{\bar{c}_t, r\}} \end{array} \right) \tag{15}$$

Then consensus of system (1) is asymptotically achieved in the sense of (2).

Proof: The proof can be done following the similar procedure as shown in theorem 1, and omitted here due to space limitation.

IV. Example

Let us reconsider example 1 for the consensus of three agents with control (4) under the sensing/communication topologies $S(t_{3k}), S(t_{3k+1})$ and $S(t_{3k+2})$ given in (5). Under the same simulation conditions, system responses are shown in 2, and consensus is reached.

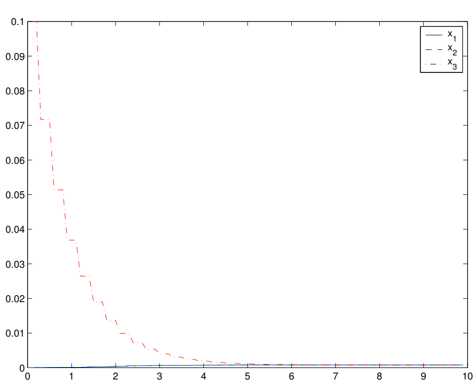

Fig. 2. System responses

V. Conclusions

In the paper, we proposed a new discontinuous cooperative control for consensus of multiagent systems with directed and switching topologies and sensing/communication delays. The proposed new design may be applied to address the cooperative control problem for truly nonlinear systems. Further research will be devoted to convergence speed and performance analysis of the proposed cooperative control.

Appendix

Proof of theorem 1: By substituting (4) into (1), we have

$$\dot{x}_i = \sum_{l=1}^{n} \alpha_{il}(s_{il}(t_k^s), x_l(t_k^s))\mathrm{sgn}(x_l(t) - x_i(t)) \triangleq \mathcal{F}_i(x), \tag{16}$$

where $i = 1, \cdots, n$ and $x = [x_1, x_2, \cdots, x_n]^T$.

We first show that if matrix sequence $\{S(t_k^s)\}$ is uniformly sequentially complete, then $x = x^*\mathbf{1}$ is the only type of equilibrium point of the closed-loop system (16) where $\mathbf{1} \in \Re^n$ a vector with all entries being 1. The proof is established by contradiction. Assume that $x^e = [x_1^e, \cdots, x_n^e]^T$ is an equilibrium point satisfying $\mathcal{F}_i(x^e) = 0, \forall i$ and $\min_i x_i^e \neq \max_i x_i^e$. Define index sets $\Phi_{\min} = \{j : x_j^e = \min_i x_i^e\}$ and $\Phi_{\max} = \{j : x_j^e = \max_i x_i^e\}$.

Since matrix sequence is uniformly sequentially complete, which is equivalent to say that the composite graph $S(t)$ has at least one globally reachable node x_g. Apparently, x_g may be in $\Phi_{\min}$ or $\Phi_{\max}$ or may not be in both sets. For any case, there must exist an index j in the compliment set of the set containing x_g while maintaining a path to x_g due to the completeness assumption. That is, $x_j \neq x_g$, which renders $\mathcal{F}_i(x) \neq 0$ for at least one i, a contradiction.

In what follows, we further show that the system (16) is Lyapunov stable, and cooperative stable (consensus can be achieved).

(a) At each time instant t, let i^* denote the index such that

$$x_{i^*}(t) = \max_j x_j(t) \tag{17}$$

we will show that $x_{i^*}(t)$ is non-increasing over time. It follows from (17) that $\mathrm{sgn}(x_l(t) - x_{i^*}(t)) \leq 0$ for all $\alpha_{i^*l} \neq 0$. Hence $\dot{x}_{i^*}(t) \leq 0$, and we conclude that the maximum value of $\max_j x_j(t)$ never increases over time. The proof of the minimum value of $\min_j x_j(t)$ never decreasing over time is similar. Lyapunov stability becomes obvious from the above conclusions.

(b) To prove consensus, we will show that the mapping defined by differential equation (16) is a contraction mapping under undirected sequentially complete network topologies. That is, we will prove that for any t, there exists a constant $\delta(t) > 0$, such that

$$\max_{i,j} \|x_i(t+\delta) - x_j(t+\delta)\| \leq \lambda \max_{i,j} \|x_i(t) - x_j(t)\|, \tag{18}$$

where $0 \leq \lambda < 1$.

Let $\Omega = \{1, \cdots, n\}$ be the set of indices on state variables, and at time t, let $x_{\max}(t) = \max_j x_j(t)$ and $x_{\min}(t) =$

$\min_j x_j(t)$. Define sub-sets $\Omega_{\max}(t)$, $\Omega_{\min}(t)$, $\Omega^c_{\max}(t) = \Omega/\Omega_{\max}(t)$, and $\Omega^c_{\min}(t) = \Omega/\Omega_{\min}(t)$ as follows:

$$\Omega_{\max}(t) = \{i^* \in \Omega : x_{i^*}(t) = x_{\max}(t)\}$$

and

$$\Omega_{\min}(t) = \{i_* \in \Omega : x_{i_*}(t) = x_{\min}(t)\}$$

To show (18), it is equivalent to prove for any t, there exists a constant $\delta(t) > 0$, such that

$$\|x_{\max}(t+\delta) - x_{\min}(t+\delta)\| \leq \lambda \|x_{\max}(t) - x_{\min}(t)\|, \quad (19)$$

for some $0 \leq \lambda < 1$.

It follows from points (a) and (b) that for any time interval $\delta(t)$, we have

$$x_{\max}(t+\delta) \leq x_{\max}(t), \quad x_{\min}(t+\delta) \geq x_{\min}(t), \quad (20)$$

thus, a weaker version of inequality (19) holds for some $0 \leq \lambda_1 \leq 1$ under arbitrary network conditions and system dynamics constraints.

Now let us show that if for any t, there exists a finite value $\delta(t)$ such that the composite network topology is complete within the time interval $[t, t+\delta)$, then (19) always holds.

Consider the evolution of $x_{i^*}(t)$ for every $i^* \in \Omega_{\max}(t)$ and $x_{i_*}(t)$ for every $i_* \in \Omega_{\min}(t)$. Several cases are in the sequel.

Case I: If there are leader nodes[†] staying in $\Omega_{\max}(t')$ for $t' \in [t, t+\delta)$, which means they don't directly or indirectly receive any information from members in $\Omega^c_{\max}(t')$ and hence $x_{\max}(t+\delta) = x_{\max}(t)$. Thus all the agents in $\Omega_{\min}(t)$ must have either direct or indirect (through agent in $\Omega_{\min}(t)$) information exchange with members in their complement set $\Omega^c_{\min}(t)$ during the time interval $[t, t+\delta(t))$, otherwise, it contradicts the completeness assumption for composite network topology within the time interval $[t, t+\delta(t))$. The question then becomes to verify $x_{\min}(t+\delta) > x_{\min}(t)$ in order to prove that (19) holds.

Now consider the evolution of $x_{i_*}(t)$ and $x_l(t)$ for $l \in \Omega^c_{\min}(t)$. According to point (a), the states $x_{i_*}(t)$ have the tendency of increase, and $x_l(t)$ have the tendency of either increase or decrease. If we can show that all $x_{i_*}(t)$ will increase in $[t, t+\delta(t))$, and the decreasing agents $x_l(t)$ will not reduce their values to some x_{i_*} at time instant t, that is, $x_{i_*}(t)$, then $x_{\min}(t+\delta) > x_{\min}(t)$ is apparent. Since by completeness assumption, every agent $i_* \in \Omega_{\min}(t)$ will have a chance to communicate with some agents in $\Omega^c_{\min}(t)$, without loss of generality, we consider that agent $i_* \in \Omega_{\min}(t)$ has the communication with agents in $\Omega^c_{\min}(t)$ right at time instant t, then we have

$$x_j(t) > x_{i_*}(t), \forall j \in \mathcal{N}_{i_*} \cap \Omega^c_{\min}(t)$$

and thus

$$\dot{x}_{i_*}(t) = \sum_{j \in \mathcal{N}_{i_*}} \alpha_{i_* j}(t) > 0. \quad (21)$$

[†]A node is called a leader node if it does not receive information from other nodes or only has communication with nodes in its current set.

It is readily seen from (21) that before network topology at time instant t switches to another topology, the value of $x_{i_*}(t+\tau)$ for $0 < \tau < \delta$ will increase within the given time interval.

On the other hand, consider the agent $l \in \Omega^c_{\min}(t)$ with decreasing speed at time instant t, and has one of its neighboring agents from $\Omega_{\min}(t)$, that is, we have

$$\begin{aligned}\dot{x}_l(t) &= \sum_{j \in \mathcal{N}_l} \alpha_{lj}(t^s_k)\mathrm{sgn}(x_j(t) - x_l(t)) < 0, \\ &\quad t \in [t^s_k, t^s_{k+1})\end{aligned} \quad (22)$$

It follows that

$$-\sum_{j \in \mathcal{N}_l} \alpha_{lj}(t^s_k) \leq \dot{x}_l(t) < 0 \quad (23)$$

and

$$x_l(t+\tau) \geq x_l(t) - \sum_{j \in \mathcal{N}_l} \alpha_{lj}(t^s_k)\tau \quad (24)$$

Note also that since $\dot{x}_l < 0$, agent l must satisfy $x_l(t^s_k) \geq \max_{j \in \mathcal{N}_l} x_j(t^s_k)$ or $\min_{j \in \mathcal{N}_l} x_j(t^s_k) < x_l(t^s_k) < \max_{j \in \mathcal{N}_l} x_j(t^s_k)$. To this end, due to the nonlinear gain $\alpha_{ij}(t^s_k)$ selected in (9) or (11), we have

$$\begin{aligned}\sum_{j \in \mathcal{N}_l} \alpha_{lj}(t^s_k) &\leq \sum_{j \in \mathcal{N}_l} \alpha_{lj}(t^s_k) \\ &< \frac{x_l(t^s_k) - \min_{j \in \mathcal{N}_l} x_j(t^s_k)}{\overline{c}_t}\end{aligned} \quad (25)$$

Note that $\min_{j \in \mathcal{N}_l} x_j(t^s_k - \tau_{ij}) \geq x_{\min}(t)$, we further have

$$\sum_{j \in \mathcal{N}_l} \alpha_{lj}(t^s_k) < \frac{x_l(t^s_k) - x_{\min}(t)}{\overline{c}_t} \quad (26)$$

together with (24), we have

$$\begin{aligned}x_l(t+\tau) &\geq x_l(t) + \frac{x_{\min}(t) - x_l(t^s_k)}{\overline{c}_t}\tau \\ &> x_{\min}(t)\end{aligned} \quad (27)$$

Therefore, we know that before network topology at time instant t switches to another topology, the value of $x_l(t+\tau)$ for some $0 < \tau < \delta$ will keep decreasing but not achieving to the minimum value at time instant t (that is, $x_{\min}(t)$) in any given time interval smaller than $\overline{c}_t$. In conclusion, for all agents $i_* \in \Omega_{\min}(t)$, their values will increase during time interval $[t, t+\delta)$; for the decreasing agents $x_l(t), l \in \Omega^c_{\min}(t)$, their values will not be able to reduce to $x_{\min}(t)$. Hence, we have $x_{\min}(t+\delta) > x_{\min}(t)$.

Case II: Similar argument is true for the case of leader nodes staying in $\Omega_{\min}(t)$ in time interval $[t, t+\delta)$. That is, for all agents $i^* \in \Omega_{\max}(t)$, their values will decrease during time interval $[t, t+\delta)$;

For agent $l \in \Omega^c_{\max}(t)$ which acquires the increasing speed at time instant t, we have

$$\begin{aligned}\dot{x}_l(t) &= \sum_{j \in \mathcal{N}_l} \alpha_{lj}(t^s_k)\mathrm{sgn}(x_j(t) - x_l(t)) > 0, \\ &\quad t \in [t^s_k, t^s_{k+1})\end{aligned} \quad (28)$$

It follows that

$$0 < \dot{x}_l(t) \leq \sum_{j\in\mathcal{N}_l} \alpha_{lj}(t_k^s) \tag{29}$$

and

$$x_l(t+\tau) \leq x_l(t) + \sum_{j\in\mathcal{N}_l} \alpha_{lj}(t_k^s)\tau \tag{30}$$

Note also that since $\dot{x}_l > 0$, agent l must satisfy $x_l(t_k^s) \leq \min_{j\in\mathcal{N}_l} x_j(t_k^s)$ or $\min_{j\in\mathcal{N}_l} x_j(t_k^s) < x_l(t_k^s) < \max_{j\in\mathcal{N}_l} x_j(t_k^s)$. Thus, according to gain selection $d_{lj}(t)$ in (10) or (11), we have

$$\begin{aligned}\sum_{j\in\mathcal{N}_l} \alpha_{lj}(t_k^s) &\leq \sum_{j\in\mathcal{N}_l} \alpha_{lj}(t_k^s) \\ &< \frac{\max_{j\in\mathcal{N}_l} x_j(t_k^s) - x_l(t_k^s)}{\overline{c}_t}\end{aligned} \tag{31}$$

which further leads to

$$\sum_{j\in\mathcal{N}_l} \alpha_{lj}(t_k^s) < \frac{x_{\max}(t) - x_l(t_k^s)}{\overline{c}_t} \tag{32}$$

and

$$x_l(t+\tau) \leq x_l(t) + \frac{x_{\max}(t) - x_l(t_k^s)}{\overline{c}_t}\tau < x_{\max}(t),$$

since $\max_{j\in\mathcal{N}_l} x_j(t_k^s) \leq x_{\max}(t)$. Thus for $x_{\min}(t+\delta) = x_{\min}(t)$, we have $x_{\max}(t+\delta) < x_{\max}(t)$.

In summary, for both **Case I** and **Case II**, inequality (19) holds.

Case III: Now we consider the case in which there are no leader nodes in $\Omega_{\max}(t)$ and $\Omega_{\min}(t)$. Then, for every $i^* \in \Omega_{\max}(t)$, we must have $x_{i^*}(t+\delta) < x_{i^*}(t)$, because system $x_{i^*}(t)$ must have state exchange with at least one element in their complement sets during time interval $[t, t+\delta)$. Same argument holds for $x_{i_*}(t)$, and we have $x_{i_*}(t+\delta) > x_{i_*}(t)$. Several sub-cases follow:

Case III-1: If $\Omega_{\max}(t) \cap \Omega_{\max}(t+\delta) \neq \emptyset$ and $\Omega_{\min}(t) \cap \Omega_{\min}(t+\delta) \neq \emptyset$, which means at least one $i^* \in \Omega_{\max}(t)$ remains staying in $\Omega_{\max}(t+\delta)$, and at least one $i_* \in \Omega_{\min}(t)$ remains staying in $\Omega_{\min}(t+\delta)$, thus (19) holds.

Case III-2: $\Omega_{\max}(t) \cap \Omega_{\max}(t+\delta) \neq \emptyset$ and $\Omega_{\min}(t) \cap \Omega_{\min}(t+\delta) = \emptyset$. It follows from $\Omega_{\max}(t) \cap \Omega_{\max}(t+\delta) \neq \emptyset$ that $x_{\max}(t+\delta) < x_{\max}(t)$. On the other hand, we have $x_{\min}(t+\delta) \geq x_{\min}(t)$ from point (b). Thus, (19) holds.

Case III-3: $\Omega_{\max}(t) \cap \Omega_{\max}(t+\delta) = \emptyset$ and $\Omega_{\min}(t) \cap \Omega_{\min}(t+\delta) \neq \emptyset$. It follows from $\Omega_{\min}(t) \cap \Omega_{\min}(t+\delta) \neq \emptyset$ that $x_{\min}(t+\delta) > x_{\min}(t)$. It then suffices to show that $x_{\max}(t+\delta) \leq x_{\max}(t)$, which is always true from point (a).

Case III-4: $\Omega_{\max}(t) \cap \Omega_{\max}(t+\delta) = \emptyset$ and $\Omega_{\min}(t) \cap \Omega_{\min}(t+\delta) = \emptyset$. This means that the entries in $\Omega_{\max}(t+\delta)$ and $\Omega_{\min}(t+\delta)$ are completely from the complement sets $\Omega^c_{\max}(t)$ and $\Omega^c_{\min}(t)$, respectively.

Due to the undirected network topology, and following the same argument in case I, we know that for any entry of $\Omega^c_{\max}(t)$, its maximum increase in $\delta(t)$ can only reach to some value less than $x_{\max}(t)$, and for any entry of $\Omega^c_{\min}(t)$, and its maximum decrease in $\delta(t)$ can only reach to some value greater than $x_{\min}(t)$. Thus, we have $x_{\max}(t+\delta) < x_{\max}(t)$ and $x_{\min}(t+\delta) > x_{\min}(t)$. This completes the proof. $\square$

REFERENCES

[1] F. Bullo, J. Cortés, and S. Martínez, *Distributed Control of Robotic Networks*, ser. Applied Mathematics Series. Princeton University Press, 2009, electronically available at http://coordinationbook.info.

[2] Y. Cao and W. Ren, "Distributed coordinated tracking with reduced interaction via a variable structure approach," *IEEE Trans. Automa. Control*, vol. 56, p. in press, 2011.

[3] Y. Cao, W. Ren, and Z. Meng, "Decentralized finite-time sliding mode estimators and their applications in decentralized finite-time formation tracking," *Systems & Control Letters*, vol. 59, pp. 522–529, 2010.

[4] G. Chen, F. L. Lewis, and L. Xie, "Finite-time distributed consensus via binary control protocols," *Automatica*, vol. 47, p. in press, 2011.

[5] D. V. Dimarogonas and K. J. Kyriakopoulos, "Connectedness preserving distributed swarm aggregation for multiple kinematic robots," *IEEE Transactions on Robotics*, vol. 24, pp. 1213–1223, 2008.

[6] J. A. Fax and R. M. Murray, "Information flow and cooperative control of vehicle formations," *IEEE Transactions on Automatic Control*, vol. 49, pp. 1465–1476, 2004.

[7] Q. Hui, "Finite-time rendezvous algorithms for mobile autonomous agents," *IEEE Trans. Automa. Control*, vol. 56, pp. 207–211, 2011.

[8] A. Jadbabaie, J. Lin, and A. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Trans. on Automatic Control*, vol. 48, pp. 988–1001, 2003.

[9] Z. Lin, M. Brouchke, and B. Francis, "Local control strategies for groups of mobile autonomous agents," *IEEE Trans. on Automatic Control*, vol. 49, pp. 622–629, 2004.

[10] Z. Lin, B. Francis, and M. Maggiore, "State agreement for continuous-time coupled nonlinear systems," *SIAM Journal on Control and Optimization*, vol. 46, pp. 288–307, 2007.

[11] L. Moreau, "Stability of multiagent systems with time-dependent communication links," *IEEE Transactions on Automatic Control*, vol. 50, pp. 169–182, 2005.

[12] B. Paden and S. Sastry, "A calculus for computing filippov's differential inclusion with appliation to the variable structure control of robot manipulators," *IEEE Transactions on Circuits and systems*, vol. 34, pp. 73–82, 1987.

[13] A. Papachristodoulou, A. Jadbabaie, and U. Munz, "Effects of delay in multi-agent consensus and oscillator synchronization," *IEEE Trans. Automa. Control*, vol. 55, pp. 1471–1477, 2010.

[14] Z. Qu, *Cooperative Control of Dynamical Systems*. London: Springer-Verlag, 2009.

[15] ——, "Cooperative control of networked nonlinear systems," in *49th IEEE Conf. on Dec. and Ctrl*, Atlanta, GA, Dec 2010, pp. 3200–3207.

[16] Z. Qu, J. Wang, and R. A. Hull, "Cooperative control of dynamical systems with application to autonomous vehicles," *IEEE Transactions on Automatic Control*, vol. 53, pp. 894–911, 2008.

[17] W. Ren and R. W. Beard, "Consensus seeking in multiagent systems under dynamically changing interaction topologies," *IEEE Transactions on Automatic Control*, vol. 50, pp. 655–661, 2005.

[18] ——, *Distributed Consensus in Multi-vehicle Cooperative Control*. London: Springer-Verlag, 2008.

[19] R. O. Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, pp. 215–233, 2007.

[20] R. O. Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Transactions on Automatic Control*, vol. 49, pp. 1520–1533, 2004.

[21] D. Shevitz and B. Paden, "Lyapunov stability theory of nonsmooth systems," *IEEE Transactions on Automatic Control*, vol. 39, pp. 1910–1994, 1994.

[22] H. G. Tanner, A. Jadbabaie, and G. J. Pappas, "Flocking in fixed and switching networks," *IEEE Transactions on Automatic Control*, vol. 52, pp. 863–868, 2007.

[23] J. Wang, Z. Qu, C.M.Ihlefeld, and R. A. Hull, "A control-design-based solution to robotic ecology: Autonomy of achieving cooperative behavior from a high-level astronaut command," *Autonomous Robots*, vol. 20, pp. 97–112, 2006.

[24] J. Wang, Z. Qu, and M. Obeng, "A distributed cooperative steering control with application to nonholonomic robots," in *49th IEEE Conf. on Dec. and Ctrl*, Atlanta, GA, Dec 2010, pp. 4571–4576.

APPROXIMATE POLICY ITERATION FOR COOPERATIVE CONTROL OF MULTIAGENT SYSTEMS UNDER LIMITED SENSING/COMMUNICATION

Jing Wang
Department of Electrical and Computer Engineering
Bradley University
Peoria, IL 61625, USA
email: jingwang@fsmail.bradley.edu

Tianyu Yang
Department of Electrical, Computer, Software, and Systems Engineering
Embry-Riddle Aeronautical University
Daytona Beach, FL 32114, USA

Gennady Staskevich, and Brian Abbe
AFRL/RISC
Rome, NY 13441, USA

ABSTRACT
In this paper, we propose an approximate policy iteration method for cooperative control of multiagent systems under the limited sensing/communication topology. By considering a class of nonlinear multiagent systems, the cooperative control problem is formulated as making all systems achieve consensus while minimizing the individual sensing/communication topology dependent cost functions. To solve the induced multiagent Hamilton-Jacobi-Bellman (HJB) equations, a multiagent policy iteration algorithm is proposed with convergence proof. Neural network parameterization is further employed to approximate value function to deal with unknown system dynamics. Through seeking the least-squares solution based on the measured online sensing/communication data, the approximate multiagent policy iteration algorithm is obtained to solve the posed optimal cooperative control problem for multi agents. Simulation results illustrate the effectiveness of the proposed optimal cooperative control.

KEY WORDS
Multiagent policy iteration, Cooperative Control, Multiagent Systems, Neural Network.

1 Introduction

Cooperative control of multiagent systems, in particular consensus control of multiagent systems, has been one of the dominating research subjects in the current control community due to numerous potential applications in the areas such as robotic network [11][14][15], power network [20], to name but a few. The research for cooperative control has been focused on two types of major issues: the necessary and sufficient multiagent network connectivity condition for information exchange [6][13], and the design of locally distributed cooperative control. While fruitful results for cooperative control design have been obtained for first-order linear systems [8][16], for second-order linear systems [21], for high-order linear systems [13], and for nonlinear systems [10][9][23], few results are available for optimal cooperative control design. There appeared some recent work in the study of optimal cooperative control, such as those in [19][2][4][12]. Nonetheless, it is still a challenge issue to systematically address the optimal cooperative control problem for more general nonlinear multiagent systems, particularly, in the presence of model uncertainties. In this paper, we develop an approximately adaptive multiagent policy iteration (MPI) algorithm to cooperatively solve the consensus problem for multiagent systems.

The result reported in this paper aims to present a dynamic programming solution to multiagent cooperative control. For multiagent optimal cooperative control, the key issue is how to establish an optimality equation and find its solution in real time. We tackle this problem by considering a general class of feedback linearizable nonlinear multiagent systems. We assume that there exist admissible cooperative controls for such kind of multiagent systems under the *complete* sensing/communication condition [13]. Since this paper is centered on the design of approximately adaptive optimal cooperative control, the fixed sensing/communication topology is imposed for ease of design. The case for more complicated time-varying sensing communication topology will be treated in future work. The optimal cooperative control problem is then formulated as making all systems achieve consensus while minimizing the individual sensing/communication topology dependent cost functions. It is shown that the optimal solution to the defined problem requires to solve a multiagent Hamilton-Jacobi-Bellman (HJB) equation. To avoid the obstacles in analytically solving multiagent HJB equation, we extend the online policy iteration approach in [18][22] to the multiagent case, and employ RBF neural networks to approximate value functions at each iteration. Through seeking the least-squares solution to estimate the optimal neural weights, a new approximately adaptive multiagent policy iteration algorithm is proposed. It is further shown that the proposed adaptive optimal cooperative control approximately solves the posed optimal consensus problem. Simulation results are provided to illustrate the effectiveness of

the proposed optimal design.

2 Problem Formulation

Consider a multiagent system which has N members and each agent assumes the general nonlinear dynamics

$$\dot{x}_i = f_i(x_i) + g_i(x_i)u_i, \tag{1}$$

where $i \in \Omega \triangleq \{1, \cdots, N\}$, $x_i(t) \in \Re^n$ is the system state, $u_i \in \Re^m$ is the control input to be designed, $f_i, g_i : \Re^n \mapsto \Re^n$ are locally Lipschitz continuous functions.

The objective of this paper is to design an optimal cooperative control $u_i(t)$ to achieve the consensus of the multiagent system (1) such that

$$\lim_{t\to\infty} x_i(t) = x^*, \quad \forall i, \tag{2}$$

while minimizing the following individual cost function for each agent i,

$$\begin{aligned} & J_i(u_i; x_i(t_0), s_{ij}x_j(t_0)) \\ = & \int_{t_0}^{\infty} \left(\sum_{j=1}^{N} (x_i - x_j)^T s_{ij} Q_{ij} (x_i - x_j) + u_i^T R_i u_i \right) dt \end{aligned} \tag{3}$$

where x^* is some constant denoting the consensus value, Q_{ij} and R_i are symmetric and positive definite matrices, and s_{ij} is a binary number describing the availability of the sensing/communication information exchange between the agent i and the agent j.

The success of solving consensus problem defined in (2) is dependent on the sensing/communication information exchange among agents, which can be described by a $N \times N$ sensing/communication matrix defined below

$$S = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1N} \\ s_{21} & s_{22} & \cdots & s_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ s_{N1} & s_{N2} & \cdots & s_{NN} \end{bmatrix}, \tag{4}$$

where $s_{ii} \equiv 1$; $s_{ij}(t) = 1$ if the ith agent can receive the information from the jth agent, and $s_{ij} = 0$ if otherwise. In general, sensing/communication matrix S could be time-varying due to the changing environment. In this paper, we focus on the design of approximately adaptive optimal cooperative control for multiagent systems under the assumption of the sensing/communication matrix S being constant and satisfying the *completeness* condition for its connectivity. The completeness condition for network connectivity was developed in [13], which can be summarized into the following definition.

Definition 2.1 *Sensing/communication matrix S is said to be complete if in every block row i of its lower triangular canonical form, there is at least one $j < i$ such that the corresponding block is nonzero.*

For more general sequentially changing sensing/communication topology, the *sequential completeness* concept of the sensing/communication matrix sequence $\{S(t)\}$ was also introduced in [13]. It is equivalent to the condition of that there exists a spanning tree in the communication graph [11], which represents the least restrictive connectivity condition for sensor/communication network in order to achieve the asymptotically cooperative consensus of the overall multiagent system. In this paper, the fixed S is considered, the completeness condition is described in definition 1. We will utilize the following assumptions for the design of optimal cooperative control.

Assumption 2.1 *The sensing/communication matrix S in (4) is complete.*

Assumption 2.2 *For nonlinear multiagent systems (1), there exist admissible cooperative control policies $u_i(t)$ to solve the consensus problem defined in (2).*

To this end, the optimal cooperative control problem can be formulated: given the nonlinear multiagent systems (1), the set of admissible cooperative control policies, and the infinite horizon cost function (3) for individual agents, find an admissible cooperative control policy such that the cost function (3) achieves its minimum.

3 The Proposed Approximate Policy Iteration for Multiagent Cooperative Control

3.1 Multiagent HJB Equation

Recall that the cost function for agent i is defined in (3), which may be rewritten as

$$\begin{aligned} & J_i(u_i; x_i(t_0), s_{ij}x_j(t_0)) \\ = & \int_{t_0}^{\infty} \left(\sum_{j\in\mathcal{N}_i} (x_i - x_j)^T Q_{ij} (x_i - x_j) + u_i^T R_i u_i \right) dt, \end{aligned} \tag{5}$$

where $\mathcal{N}_i = \{j \in \Omega | s_{ij} \neq 0\}$ denotes the neighbor set of agent i. The following lemma is instrumental in developing the multiagent Hamilton-Jacobi-Bellman (HJB) equation.

Lemma 3.1 *For admissible cooperative control $u_i(t)$, if there exists a positive definite continuously differentiable function $V_i(x_i, s_{ij}x_j; u_i)$ satisfying the following property*

$$\begin{aligned} & \frac{\partial V_i^T}{\partial x_i}(f_i(x_i) + g_i(x_i)u_i) \\ & + \sum_{j\in\mathcal{N}_j} \frac{\partial V_i^T}{\partial x_j}(f_i(x_j) + g_i(x_j)u_j) \\ & + \sum_{j\in\mathcal{N}_i} (x_j - x_i)^T Q_{ij}(x_j - x_i) + u_i^T R_i u_i = 0 \end{aligned} \tag{6}$$

and the boundary condition $V_i(x_i(\infty), s_{ij}x_j(\infty); u_i) = 0$, then $V_i(x_i, s_{ij}x_j; u_i)$ is the value function for system (1) for all t, and

$$V_i(x_i(t_0), s_{ij}x_j(t_0); u_i) = J_i(u_i; x_i(t_0), s_{ij}x_j(t_0)) \tag{7}$$

To this end, it follows from lemma 3.1 and Bellman's principle of optimality [3], we know that the optimal value function $V_i^*(x_i(t), s_{ij}x_j(t))$ approximately satisfies for small $\Delta \to 0$

$$\begin{aligned} & V_i^*(x_i(t), s_{ij}x_j(t)) \\ & \simeq \min_{u_i}[l(x_i(t), x_j(t), u_i)\Delta + V_i^*(x_i(t+\Delta), s_{ij}x_j(t+\Delta)], \end{aligned} \tag{8}$$

where $l(x_i(t), x_j(t), u_i) \overset{\triangle}{=} \sum_{j\in\mathcal{N}_i}(x_i - x_j)^T Q_{ij}(x_i - x_j) + u_i^T R_i u_i$, $x_i(t+\Delta) \simeq x_i(t) + (f_i(x_i) + g_i(x_i)u_i)\Delta$, and $x_j(t+\Delta) \simeq x_j(t) + (f_i(x_j) + g_i(x_j)u_j)\Delta$. Since V_i^* is continuously differentiable, we may write (for $\Delta \to 0$)

$$\begin{aligned} V_i^*(x_i(t+\Delta), s_{ij}x_j(t+\Delta)) &\simeq V_i^*(x_i(t), s_{ij}x_j(t)) \\ &+\frac{\partial V_i^{*T}}{\partial x_i}(x_i(t), x_j(t))[f_i(x_i) + g_i(x_i)u_i]\Delta \\ &+\sum_{j\in\mathcal{N}_i}\frac{\partial V_i^{*T}}{\partial x_j}(x_i(t), x_j(t))[f_i(x_j) + g_i(x_j)u_j]\Delta. \end{aligned} \tag{9}$$

Substituting (9) in (8) we obtain the multiagent HJB equation

$$0 = \min_{u_i} H_i(x_i, s_{ij}x_j, u_i, V_i^*) \tag{10}$$

where the multiagent Hamiltonian is defined as

$$\begin{aligned} & H_i(x_i, s_{ij}x_j, u_i, V_i) \\ &= \sum_{j\in\mathcal{N}_i}(x_j - x_i)^T Q_{ij}(x_j - x_i) + u_i^T R_i u_i \\ &\quad+\frac{\partial V_i^T}{\partial x_i}(f_i(x_i) + g_i(x_i)u_i) \\ &\quad+\sum_{j\in\mathcal{N}_i}\frac{\partial V_i^T}{\partial x_j}(f_i(x_j) + g_i(x_j)u_j) \end{aligned} \tag{11}$$

The minimum with respect to u_i is obtained by solving $\frac{\partial H_i(x_i, s_{ij}x_j, u_i, V_i^*)}{\partial u_i} = 0$, that is,

$$2u_i^T R_i + \frac{\partial V_i^{*T}}{\partial x_i} g_i(x_i) = 0 \tag{12}$$

yielding the optimal cooperative control

$$u_i^* = -\frac{1}{2}R_i^{-1} g_i^T \frac{\partial V_i^*}{\partial x_i} \tag{13}$$

Substituting (13) into (10) yields

$$\begin{aligned} 0 &= \sum_{j\in\mathcal{N}_i}(x_j - x_i)^T Q_{ij}(x_j - x_i) + \frac{\partial V_i^{*T}}{\partial x_i} f_i(x_i) \\ &\quad-\frac{1}{4}\frac{\partial V_i^{*T}}{\partial x_i} g_i(x_i) R_i^{-1} g_i(x_i)^T \frac{\partial V_i^*}{\partial x_i} \\ &\quad+\sum_{j\in\mathcal{N}_i}\frac{\partial V_i^{*T}}{\partial x_i}(f_i(x_j) + g_i(x_j)u_j), \end{aligned} \tag{14}$$

with the associated boundary condition $V_i^*(x_i^*, s_{ij}x_j^*) = 0$, which requires that the optimal value must be null when evaluated on an extremal trajectory (all agents in the set $\{i, \mathcal{N}_i\}$ reach consensus.)

Equation (14) is the *mutliagent HJB equation*. Its solution would provide the optimal cooperative control in (13). However, it is difficult to solve mainly for two reasons. First, equation (14) is a nonlinear partial differential equation, and it is in general impossible to solve this equation in analytic form. Second, the coupling terms $\sum_{j\in\mathcal{N}_i}\frac{\partial V_i^{*T}}{\partial x_i}(f_i(x_j) + g_i(x_j)u_j)$ cause extra difficulty due to involvement of u_j which may require information propagation from agents not in the neighboring set $\mathcal{N}_i$.

3.2 Multiagent Policy Iteration Algorithm

Motivated by the policy iteration algorithm for solving HJB equation for single agent systems in [18], in what follows, we provide the mutliagent policy iteration algorithm for approximately solving the multiagent HJB equation (14). The proposed multiagent policy iteration algorithm consists of the following two steps:

Step 1: Policy evaluation. Find an admissible cooperative control policy $u_{i,0}(x_i, s_{ij}x_j)$. For any integer $l \geq 0$ denoting the iteration index, solve for $V_{i,l}(x_i, s_{ij}x_j; u_{i,l})$ using

$$\begin{aligned} 0 &= \sum_{j\in\mathcal{N}_i}(x_j - x_i)^T Q_{ij}(x_j - x_i) + u_{i,l}^T R_i u_{i,l} \\ &\quad+\frac{\partial V_{i,l}^T}{\partial x_i}(f_i(x_i) + g_i(x_i)u_{i,l}) \\ &\quad+\sum_{j\in\mathcal{N}_i}\frac{\partial V_{i,l}^T}{\partial x_j}(f_i(x_j) + g_i(x_j)u_j), \end{aligned} \tag{15}$$

with $V_{i,l}(x^*, s_{ij}x^*) = 0$.

Step 2: Policy improvement. Update the control policy by

$$u_{i,l+1} = -\frac{1}{2}R_i^{-1} g_i^T \frac{\partial V_{i,l}}{\partial x_i} \tag{16}$$

The convergence of the multiagent policy iteration algorithm given in (15) and (16) is summarized into the following theorem.

Theorem 1 *If a sequence of pairs $\{V_{i,l}, u_{i,l+1}\}$ is generated by (15) and (16), then the corresponding value function $V_{i,l}$ satisfying*

$$V_{i,l+1} \leq V_{i,l} \tag{17}$$

and

$$\lim_{l\to\infty} V_{i,l} = V_i^* \tag{18}$$

The proof of theorem 1 can be done following the similar lines of reasoning as that of theorem 4 in [18]. Details are omitted due to space limitation. The proposed multiagent policy iteration algorithm relieves the nonlinearity obstacle in solving multiagent HJB equation (14) to a certain level in the sense of only dealing with a linear partial differential equation. For instance, by linearly parameterizing $\frac{\partial V_{i,l}^T}{\partial x_i}$, the solution to (15) can be obtained from a set of linear algebra equations. However, $u_{j,l}$ for $j \in \mathcal{N}_i$ are still needed in solving (15), which might be hard to directly be sensed and/or communicated.

To avoid this obstacle, we note that the solution $V_{i,l}$ to (15) is actually the value function for system (1) at the iteration l, since it satisfies the properties in lemma 3.1. Thus, we obtain

$$V_{i,l}(x_i(t), s_{ij}x_j(t); u_{i,l}) = J_i(u_{i,l}; x_i(t), s_{ij}x_j(t)), \forall t.$$

It follows from the above equation and (3) that

$$\begin{aligned} & V_{i,l}(x_i(t), s_{ij}x_j(t); u_{i,l}(t)) \\ &= \int_t^{t+T} \left(\sum_{j\in\mathcal{N}_i} (x_j - x_i)^T Q_{ij}(x_j - x_i) + u_{i,l}^T R_i u_{i,l} \right) dt \\ &+ V_{i,l}(x_i(t+T), s_{ij}x_j(t+T); u_{i,l}(t+T)), \end{aligned} \tag{19}$$

where $T > 0$ represents certain time interval. To this end, the **policy evaluation** step in the proposed multiagent policy iteration algorithm can be replaced by equation (19) for solving for $V_{i,l}$ based on the available information $x_i(t), x_j(t)$ and $u_{i,l}(t)$ during the time interval $[t, t+T]$.

3.3 Approximate Policy Iteration

A significant advantage of the proposed multiagent policy iteration algorithm is that it iteratively generates a sequence of pairs $\{V_{i,l}, u_{i,l+1}\}$ through (19) and (16) by only using the available local information x_i, x_j and u_i for agent i, which monotonically converges to the optimal value V_i^* and u_i^*. It is apparent that the key is to solve for $V_{i,l}$ from (19). To facilitate the design and for the ease of implementation, in the sequel, we hypothesize that $V_{i,l}$ has a linearly parameterized form as

$$V_{i,l}(x_i, s_{ij}x_j) = \sum_{\mu=1}^{m} \phi_{i_\mu,l}(x_i, s_{ij}x_j)\theta_{i_\mu,l}^* = \Phi_{i,l}^T \theta_{i,l}^*, \tag{20}$$

where $\Phi_{i,l} = [\phi_{i_1,l}, \phi_{i_2,l}, \cdots, \phi_{i_m,l}]^T \in \Re^m$ are some known basis functions, and $\theta_{i,l}^* = [\theta_{i_1,l}, \theta_{i_2,l}, \cdots, \theta_{i_m,l}]^T \in \Re^m$ are unknown constant parameters to be estimated.

It is worth pointing out that the value functions $V_{i,l}$ are generally nonlinear and may not assume the exact parametric form as that in (20). In that sense, a linearly parameterized approximator can be used to approximate unknown nonlinear value function $V_{i,l}$. Several function approximators are available for this purpose, such as, radial basis function (RBF) neural networks [5, 17], high-order neural networks [7] and fuzzy systems [24], which are described as $W^T S(z)$ with input vector $z \in R^n$, weight vector $W \in R^l$, node number l, and basis function vector $S(z) \in R^l$. Universal approximation results indicate that, if l is chosen sufficiently large, then $W^T S(z)$ can approximate any continuous function to any desired accuracy over a compact set [7, 17].

In this paper, we assume that the value functions $V_{i,l}$ are approximated by RBF neural networks. That is, for the unknown value functions $V_{i,l}(x_i, s_{ij}x_j)$, we have the following approximation over some compact set Ω_i

$$V_{i,l}(x_i, s_{ij}x_j) = \Phi_{i,l}^T(\bar{x}_i)\theta_{i,l}^* + \omega_{i,l}(\bar{x}_i), \ \ \forall \bar{x}_i \in \Omega_i \tag{21}$$

where $\bar{x}_i = [s_{i1}x_1, s_{i2}x_2, \cdots, x_i, \cdots, s_{ij}x_j, \cdots, s_{iN}x_N]^T$, $\theta_{i,l}^* \in R^{l_i}$ is an unknown constant parameter vector, the neural network node number $l_i > 1$, $\omega_{i,l}(\bar{x}_i)$ is the approximation error, and $\Phi_{i,l}(\bar{x}_i) = [\phi_{i_1,l}, \phi_{i_2,l}, \cdots, \phi_{i_{l_i},l}]^T$ is the known basis function vector.

Upon using the function approximator (21), the policy evaluation equation in (19) becomes

$$\begin{aligned} & \int_t^{t+T} \left(\sum_{j\in\mathcal{N}_i} (x_j - x_i)^T Q_{ij}(x_j - x_i) + u_{i,l}^T R_i u_{i,l} \right) dt \\ &= [\Phi_{i,l}(\bar{x}_i(t)) - \Phi_{i,l}(\bar{x}_i(t+T))]^T \theta_{i,l}^* + \bar{\omega}_{i,l}(t), \end{aligned} \tag{22}$$

where $\bar{\omega}_{i,l}(t) = \omega_{i,l}(t) - \omega_{i,l}(t+T)$.

Remark 3.1 *Based on the universal approximation theorem [7, 17], approximation error $\omega_{i,l}(\bar{x}_i)$ will uniformly converge to zero as the neural network node number $l_i \to \infty$. In other words, $|V_{i,l} - \Phi_{i,l}^T\theta_{i,l}^*| \to 0$ as $l_i \to \infty$. Thus, $\bar{\omega}_{i,l}(t) \to 0$ as $l_i \to \infty$, which implies that (22) can be used as an approximation for the policy evaluation in the proposed multiagent policy iteration algorithm.* ◇

It follows from (22) that $\theta_{i,l}^*$ provides the best approximate solution for the policy evaluation. However, its value is unknown, and needs to identified online. Let $\theta_{i,l}(t)$ be the estimate of $\theta_{i,l}^*$, and equation (22) becomes

$$\begin{aligned} & \int_t^{t+T} \left(\sum_{j\in\mathcal{N}_i} (x_j - x_i)^T Q_{ij}(x_j - x_i) + u_{i,l}^T R_i u_{i,l} \right) dt \\ &= [\Phi_{i,l}(\bar{x}_i(t)) - \Phi_{i,l}(\bar{x}_i(t+T))]^T \theta_{i,l}(t) + e_{i,l}(t), \end{aligned} \tag{23}$$

where $e_{i,l}(t) = [\Phi_{i,l}(\bar{x}_i(t)) - \Phi_{i,l}(\bar{x}_i(t+T))]^T \tilde{\theta}_{i,l}(t) + \bar{\omega}_{i,l}(t)$, $\tilde{\theta}_{i,l}(t) = \theta_{i,l}^* - \theta_{i,l}(t)$. Thus, given any admissible cooperative control, the parameter $\theta_{i,l}$ should be chosen to minimize the squared approximation residual error $e_{i,l}^2(t)$. As $\theta_{i,l}(t) \to \theta_{i,l}^*$, it is obvious that $e_{i,l}(t) \to \bar{\omega}_{i,l}$.

In what follows, we present the proposed adaptive law for $\theta_{i,l}$ using the least-squares estimation. To proceed with the proposed adaptive design, we introduce an infinite sequence of time instants $\{t_k \triangleq t_0 + kT\}$ for $k \in \mathcal{N} \triangleq \{0, 1, \cdots, \}$ with $T > 0$ the sampling time. The proposed adaptive multiagent policy iteration algorithm relies on two types of updating intervals:

1) Control action interval $[t_k, t_{k+1})$, in which the same control policy $u_{i,l}$ will be applied;

2) Observation interval $[t_k, t_{k+n})$, in which n sets of control action with the same control policy $u_{i,l}$ will be applied, and observation data during the n intervals will be used for the least-squares estimation of $\theta_{i,l}$.

Remark 3.2 *Note that the observation interval is imposed for the least-squares solution of $V_{i,l}$ based on (23) for the* **policy evaluation** *step in the proposed multiagent policy iteration algorithm, while the control action interval corresponds to the implementation of control $u_{i,l}$ from the* **policy improvement** *step.* ◇

For notational convenience, let us define

$$z_i(t_k) = \int_{t_k}^{t_{k+1}} \left(\sum_{j\in\mathcal{N}_i} (x_j - x_i)^T Q_{ij}(x_j - x_i) + u_{i,l}^T R_i u_{i,l} \right) dt$$

and

$$\Psi_{i,l}(t_k) = \Phi_{i,l}(\bar{x}_i(t_k)) - \Phi_{i,l}(\bar{x}_i(t_{k+1})).$$

Substituting this into (23) yields

$$z_i(t_k) = \Psi_{i,l}(t_k)^T \theta_{i,l} + e_{i,l}(t_k) \tag{24}$$

The model in (24) is the *regression model* for policy iteration and $\Psi_{i,l}$ is called the *regressor*. Through the observation interval $[t_k, t_{k+n}]$, pairs of observations and regressors $\{(z_i(t_{k+\mu}), \Psi_{i,l}(t_{k+\mu})), \mu = 0, 1, n-1\}$ are obtained by using control policy $u_{i,l}$. The parameter $\theta_{i,l}$ will be chosen to minimize the least-squares loss function

$$L(\theta_{i,l}, t_k) = \frac{1}{2}\sum_{\mu=1}^{n-1} \left(z_i(t_{k+\mu}) - \Psi_{i,l}(t_{k+\mu})^T \theta_{i,l} \right)^2.$$

To this end, standard least-squares estimation algorithm renders

$$\theta_{i,l} = \left(\Xi_{i,l}^T \Xi_{i,l}\right)^{-1} \Xi_{i,l}^T Z_{i,l} \tag{25}$$

where $Z_{i,l} = [z_i(t_k), z_i(t_{k+1}), \cdots, z_i(t_{k+n-1})]^T$, and $\Xi_{i,l} = [\Psi_{i,l}^T(t_k), \cdots, \Psi_{i,l}^T(t_{k+n-1})]^T$. Thus, according to **policy improvement** step in (16), and noting $\frac{\partial V_{i,l}}{\partial x_i} = \frac{\partial \Phi_{i,l}^T}{\partial x_i}\theta_{i,l}$ the control law is

$$u_{i,l+1} = -\frac{1}{2} R_i^{-1} g_i^T \frac{\partial \Phi_{i,l}^T}{\partial x_i} \theta_{i,l}. \tag{26}$$

The above results can be summarized into the following proposition.

Proposition 3.1 *Under assumptions 2.1, 2.2 and* **??**, *the control law (26) with adaptive law (25) approximately solves the optimal cooperative consensus problem for multiagent nonlinear system (1) by minimizing the cost function (3).*

Proof: The proof can naturally be done following the above multiagent policy iteration design steps, the least-squares estimation, the claims in lemma 3.1, theorem 1 as well as the universal approximation theorem for neural network function approximation. □

Remark 3.3 *The implementation of estimation algorithm in (25) requires an excitation condition for matrix $\Xi_{i,l}^T \Xi_{i,l}$, which could be satisfied with the careful choices of basis function for neural network approximators. To further reduce the computation load due to the computation requirement for matrix inverse, in what follows, we give a simplified adaptive recursive algorithm for $\theta_{i,l}$.* ◇

The simplified adaptive recursive update for $\theta_{i,l}$ is based on Kaczmarz's project algorithm [1]. That is, one pair of data $\{z_i(t_k), \Psi_{i,l}(t_k)\}$ generates an estimate $\theta_{i,l}(t_k)$. Once a new measurement pair $\{z_i(t_{k+1}), \Psi_{i,l}(t_{k+1})\}$ is obtained, it is natural to choose the new estimate $\theta_{i,l}(t_{k+1})$ as that minimizes the following cost function

$$\begin{aligned} L &= \frac{1}{2}(\theta_{i,l}(t_{k+1}) - \theta_{i,l}(t_k))^T(\theta_{i,l}(t_{k+1}) - \theta_{i,l}(t_k)) \\ &\quad + \lambda(z_i(t_{k+1}) - \Psi_{i,l}(t_{k+1})\theta_{i,l}(t_{k+1})), \end{aligned} \tag{27}$$

where λ is a Lagrangian multiplier. Taking derivatives with respect to $\theta_{i,l}(t_{k+1})$ and λ, we obtain

$$\theta_{i,l}(t_{k+1}) - \theta_{i,l}(t_k) - \lambda\Psi_{i,l}(t_{k+1}) = 0, \tag{28}$$

$$z_i(t_{k+1}) - \Psi_{i,l}(t_{k+1})\theta_{i,l}(t_{k+1}) = 0. \tag{29}$$

Solving the above equations yields

$$\begin{aligned} \theta_{i,l}(t_{k+1}) &= \theta_{i,l}(t_k) \\ &+ \frac{\Psi_{i,l}(t_{k+1})}{\Psi_{i,l}(t_{k+1})^T \Psi_{i,l}(t_{k+1})}(z_i(t_{k+1}) - \Psi_{i,l}(t_{k+1})\theta_{i,l}(t_k)) \end{aligned} \tag{30}$$

To avoid the possible singularity for the term $\Psi_{i,l}(t_{k+1})^T \Psi_{i,l}(t_{k+1})$ for computation stability, a modified algorithm for (30) would be used in practice as given below by the double-column formula (31), where $\gamma > 0$ is the learning rate and α is a positive constant.

In summary, the proposed approximately adaptive multiagent policy iteration (MPI) algorithm is given in Algorithm 1.

Algorithm 1 Approximately Adaptive MPI Algorithm

1: Let $l = 0$. Given initial states $x_i(t_0)$, $s_{ij}x_j(t_0)$, let the initial admissible cooperative control policy be $u_{i,0}$.
2: Employ the control policy $u_{i,l}$, and within the observation interval $[t_{l\times n}, t_{(l+1)n-1}]$, collect the data pairs
$$\{(z_i(t_{l\times n+\mu}), \Psi_{i,0}(t_{l\times n+\mu})), \mu = 0, 1, n-1\}$$
3: Solve for $\theta_{i,l}$ from (31).
4: Solve for $u_{i,l+1}$ from (26).
5: Let $l \leftarrow l + 1$, and go to step 2, until
$$\|\theta_{i,l+1} - \theta_{i,l}\|^2 \le \epsilon,$$
where $\epsilon > 0$ is a sufficiently small predefined threshold.

$$\theta_{i,l}(t_{k+1}) = \theta_{i,l}(t_k) + \frac{\gamma \Psi_{i,l}(t_{k+1})}{\Psi_{i,l}(t_{k+1})^T \Psi_{i,l}(t_{k+1}) + \alpha}(z_i(t_{k+1}) - \Psi_{i,l}(t_{k+1})\theta_{i,l}(t_k)) \tag{31}$$

4 Simulation Results

To illustrate the proposed approximately adaptive cooperative control, we consider a simple multiagent system with 3 agents modeled by the following single integrators

$$\dot{x}_i = u_i, \quad i = 1, 2, 3 \tag{32}$$

where $x_i \in \Re$, and $u_i \in \Re$. Let the sensing/communication topology among 3 agents be given by

$$S = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

Apparently, S matrix is complete, and admissible cooperative control exists for the consensus of (32). Select the weight matrices in (3) as $Q_{ij} = 1, R_i = 0.25$ for simulation purpose. We use a single neural node approximator for each value function $V_{i,l}$. Based on S matrix, we choose the basic functions as $\Phi_{1,l} = (x_1 - x_2)^2$, $\Phi_{2,l} = (x_2 - x_3)^2$ and $\Phi_{3,l} = (x_3 - x_1)^2$ for value functions $V_{1,l}, V_{2,l}$ and $V_{3,l}$, respectively. System initial states are $x_1(0) = 0.5, x_2(0) = 0.2$ and $x_3(0) = 0.3$. Applying the proposed approximately adaptive MPI algorithm in Algorithm 1, the correspondingly cooperative controls are of the form (26), and the adaptive laws for $\theta_{i,l}$ are given by (31). Figure 1 shows that system states consensus is achieved, figure 2 displays the instantaneous cost values. Figure 3 illustrates the optimal cooperative control inputs, and the convergence of neural network weights estimates is shown in figure 4.

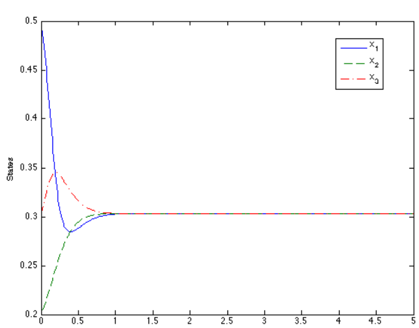

Figure 1. Consensus of $x_i(t)$

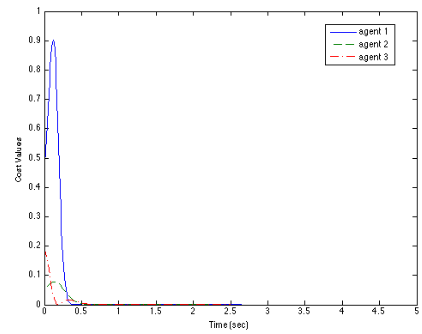

Figure 2. Instantaneous cost values versus time

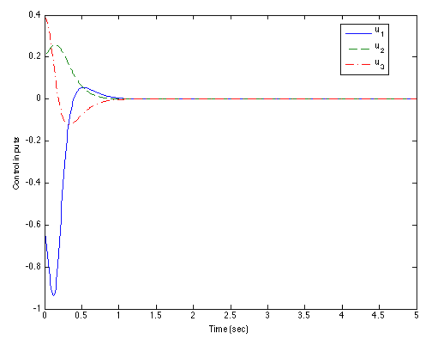

Figure 3. Optimal cooperative controls

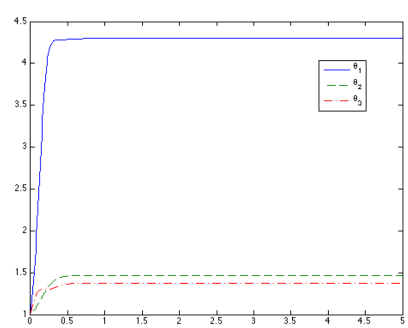

Figure 4. Parameters of neural networks versus time

5 Conclusions

In this paper, we proposed a new approximately adaptive online multiagent policy iteration algorithm to address the

optimal cooperative consensus control problem for a general class of nonlinear multiagent systems. The proposed design relies on iterative policy evaluation and policy improvement by using neural network based online adaptive estimation for optimal value functions. Simulation results further verified the effectiveness of the proposed design.

Acknowledgement

This work was supported by the Air Force Research Laboratory under grant FA8750-13-1-0109.

References

[1] K. J. Astrom and B. Witenmark. *Adaptive Control*. Addison-Wesley, Reading, MA, 1995.

[2] D. Bauso, L. Giarre, and R. Pesenti. Nonlinear protocols for optimal distributed consensus in networks of dynamic agents. *Systems and Control Letters*, 55:918–928, 2006.

[3] R. E. Bellman and S. E. Dreyfus. *Applied Dynamic Programming*. Princeton University Press, Princeton, NJ, 1962.

[4] Y. Cao and W. Ren. Optimal linear consensus algorithms: an lqr perspective. *IEEE Trans. on System, Man and Cybernetics Part B: Cybernetics*, 40:819–830, 2010.

[5] S. S. Ge, C. C. Hang, T. H. Lee, and T. Zhang. *Stable Adaptive Neural Network Control*. Kluwer Academic Publisher, Norwell, USA, 2001.

[6] A. Jadbabaie, J. Lin, and A.S. Morse. Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Trans. on Automatic Control*, 48:988–1001, 2003.

[7] E. B. Kosmatopoulos, M. M. Polycarpou, M. A. Christodoulou, and P. A. Ioannou. High-order neural network structures for identification of dynamical systems. *IEEE Trans. Neural Networks*, 6:422–431, 1995.

[8] Z. Lin, M. Brouchke, and B. Francis. Local control strategies for groups of mobile autonomous agents. *IEEE Trans. on Automatic Control*, 49:622–629, 2004.

[9] Z. Lin, B. Francis, and M. Maggiore. State agreement for continuous-time coupled nonlinear systems. *SIAM Journal on Control and Optimization*, 46:288–307, 2007.

[10] L. Moreau. Stability of multiagent systems with time-dependent communication links. *IEEE Transactions on Automatic Control*, 50:169–182, 2005.

[11] Z. Qu. *Cooperative Control of Dynamical Systems*. Springer-Verlag, London, 2009.

[12] Z. Qu and M. Simaan. An analytic solution to the optimal design of information structure and cooperative control in networked systems. In *51th IEEE Conf. on Dec. and Ctrl*, pages 4015–4022, Maui, HI, Dec 2012.

[13] Z. Qu, J. Wang, and R. A. Hull. Cooperative control of dynamical systems with application to autonomous vehicles. *IEEE Transactions on Automatic Control*, 53:894–911, 2008.

[14] W. Ren and R. W. Beard. *Distributed Consensus in Multi-vehicle Cooperative Control*. Springer-Verlag, London, 2008.

[15] R. O. Saber, J. Alex Fax, and R. M. Murray. Consensus and cooperation in networked multi-agent systems. *Proceedings of the IEEE*, 95:215–233, 2007.

[16] R. O. Saber and R. M. Murray. Consensus problems in networks of agents with switching topology and time-delays. *IEEE Transactions on Automatic Control*, 49:1520–1533, 2004.

[17] R. M. Sanner and J. E. Slotine. Gaussian networks for direct adaptive control. *IEEE Trans. Neural Networks*, 3:837–863, 1992.

[18] G. N. Saridis and C. G. Lee. An approximation theory of optimal control for trainable manipulators. *IEEE Trans. on Systems, Man and Cybernetics*, 9:152–159, 1979.

[19] E. Semsar-Kazerooni and K. Khorasani. Optimal consensus algorithms for cooperative team of agents subject to partial information. *Automatica*, 44:2766–2777, 2008.

[20] D. D. Siljak. *Large-scale dynamic systems: stability and structure*. North-Holland, New York, 1978.

[21] H. G. Tanner, A. Jadbabaie, and G. J. Pappas. Flocking in fixed and switching networks. *IEEE Transactions on Automatic Control*, 52:863–868, 2007.

[22] D. Vrabie and F. L. Lewis. Neural network approach to continuous-time direct adaptive optimal control for partially-unknown nonlinear systems. *Neural Networks*, 22:237–246, 2009.

[23] J. Wang, M. Obeng, Z. Qu, T. Yang, G. Staskevich, and B. Abbe. Discontinuous cooperative control for consensus of multiagent systems with switching topologies and time-delays. In *52th IEEE Conf. on Dec. and Ctrl*, pages 7333–7338, Florence, Italy, Dec 2013.

[24] L. X. Wang. *Adaptive Fuzzy Systems and Control: Design and Analysis*. Prentice-Hall, Englewood Cliffs, NJ, 1994.

Experimental Validation of Distributed Cooperative Control for Mobile Agents with Switching Topologies and Time-Delays*

Junzhen Shao[1], Jing Wang[2] and Tianyu Yang[1]

***Abstract*— In this paper, we present practical experimental results to demonstrate a control law for consensus of multiagent systems with switching topologies and time delays. The nonlinear control law utilizes nonlinear cooperative control gains and uses contraction mapping to achieve consensus of multiagent systems. The testing platform we used consists of a number of mobile robots. We present the effectiveness of the control law design by Aria mobile robots with applications in distributed cooperative formation control. Computer simulations and hardware experiments presented include point consensus control and formation control, both with changing topologies and time-delays using directed and undirected communication topologies.**

I. INTRODUCTION

Cooperative control[5][10] aims at achieving consensus or agreement dynamics in a multiagent system. It is an area of research lying at the intersection of systems and graph theory. A prominent application area of cooperative control is autonomous systems, especially for military and government applications. The development of single agent systems is increasingly mature in recent years. For example, unmanned aerial vehicles (UAV) and autonomous underwater vehicles (AUV) play an important role in applications in severe environments or classified operations. On the other hand Cooperative control of multiagent systems can enhance the system performance for applications like patrolling, monitoring, etc. The design of cooperative control is closely related to system dynamics. For linear systems, the dynamics can be simplified to the first-order integrator model or the double integrator model[1][2]. For nonlinear systems, which are more relevant to real world applications, cooperative control becomes much more complicated, and large gaps exist between theoretical system design and practical applications[1][3].

The multigent system is a computerized system of multiple interacting intelligent agents within an environment[5][13], and the agents work together to accomplish certain tasks. Each agent in the system has the capability of self-operating. There are two key topics in the research of multiagent systems: the design of cooperative control laws, and the controllability of networks. The network communication topology plays a key role in accomplishing consensus tasks. From this perspective, several different communication strategies have been proposed[2][7]. A popular method is the leader-follower model, in which one agent plays as the leader, and other agents communicate with the leader when performing the tasks. This model has little communication requirement and short reaction time. Nevertheless, the entire system breaks down once the leader agent is disabled.

In this paper, we experimentally validate the effectiveness of the nonlinear cooperative control proposed in [1], which is demonstrated through discontinuous cooperative control for consensus of multiagent systems with switching topologies and time-delays using mobile robots[1][2][12][15]. By designing nonlinear piecewise control gains, the consensus or formation of multiangent systems with switching topologies and time-delays are achieved both in software simulations and hardware experiments.

II. PROBLEM FORMULATION

The dynamics of a group of mobile agents are expressed as

$$\dot{x}_i = v_i \cos\theta_i, \dot{y}_i = v_i \sin\theta_i, \dot{\theta}_i = \omega_i \tag{1}$$

where x_i and y_i denote the position of the ith agent, θ_i shows the orientation which is based on the driving velocity v_i and steering velocity ω_i. In this case $(x_i, y_i) \in \mathbb{R}^2$, $(v_i, \omega_i) \in \mathbb{R}$ and $i \in 1, \cdots, n$.

Let us define the desired trajectory for the group of agents as

$$q_0(t) = [x_0(t), y_0(t)]^T \in \mathbb{R}^2 \tag{2}$$

And the motion frame is denoted as $\mathbb{F}(t)$, which can be considered as a constraint in geometric coordinates in terms of relative positions of the robots. $\mathbb{F}(t)$ consists of $q_0(t)$ and the orthonormal vectors, e_1, e_2, as defined below,

$$e_1(t) = \begin{bmatrix} e_{11}(t) \\ e_{12}(t) \end{bmatrix} = \begin{bmatrix} \frac{\dot{x}_0(t)}{\sqrt{[\dot{x}_0(t)]^2+[\dot{y}_0(t)]^2}} \\ \frac{\dot{y}_0(t)}{\sqrt{[\dot{x}_0(t)]^2+[\dot{y}_0(t)]^2}} \end{bmatrix}, \tag{3}$$

$$e_2(t) = \begin{bmatrix} e_{21}(t) \\ e_{22}(t) \end{bmatrix} = \begin{bmatrix} \frac{\dot{x}_0(t)}{\sqrt{[\dot{x}_0(t)]^2+[\dot{y}_0(t)]^2}} \\ \frac{\dot{y}_0(t)}{\sqrt{[\dot{x}_0(t)]^2+[\dot{y}_0(t)]^2}} \end{bmatrix}, \tag{4}$$

Based on the orthonormal vectors $e_i(t)$ and the trajectory information $q_0(t)$, the agent position is given as,

$$P_i(t) = \alpha_{i1} e_1(t) + \alpha_{i2} e_2(t), \tag{5}$$

*This work was supported by Air Force Research Laboratory FA8750-13-1-0109, Bethune-Cookman University and Embry-Riddle Aeronautical University

[1]Tianyu Yang and Junzhen Shao are with the department of Electrical Engineering and Computer Science, Embry-Riddle Aeronautical University, Daytona Beach, FL 32114, USA sz030247@hotmail.com

[2]Jing Wang is with Faculty of Computer Engineering, School of Science, Engineering and Mathematics, Bethune-Cookman University, Daytona Beach, FL 32114, USA

978-1-4799-4774-4/14/$31.00 ©2014 IEEE

where $P_i(t)$ is the position of the ith robot, and α_{ij} are constants determining the formation shape.

The sensing/communication information exchange among the fleet agents can be expressed by the sensing/communication matrix,

$$S(t_k^s) = \begin{bmatrix} s_{11} & a_{12}(t_k^s) & \cdots & a_{1m}(t_k^s) \\ s_{21}(t_k^s) & a_{22} & \cdots & a_{2m}(t_k^s) \\ \vdots & \vdots & \ddots & \vdots \\ s_{m1}(t_k^s) & s_{m2}(t_k^s) & \cdots & a_{mm} \end{bmatrix}, \tag{6}$$

where at $(t_k^s) : k = 0, 1, \cdots$, the ith agent receives velocity, orientation and position information from agent j, if $s_{ij}(t_k^s) = 1$. Otherwise, if $s_{ij}(t_k^s) = 0$, there is no communication between agent i and agent j.

Define the $sign(z)$ function as,

$$sign(z) = \begin{cases} 1, & z > 0 \\ 0, & z = 0 \\ -1, & z < 0 \end{cases} \tag{7}$$

For $s_{ij}(t_k^s) \neq 0$, the control model can be expressed as,

$$u_i(t) = \sum_{j=1}^{n} \alpha(s_{ij}(t_k^s), P_j(t_k^s)) sign(P_j(t) - P_i(t)), \tag{8}$$

where $t \in [t_k^s, t_{k+1}^s]$, and $\alpha(,)$ is a nonlinear control gain.

III. Control Design

In this paper we use formation control to demonstrate the consensus of multiagent systems with the new control law (8). First, we use the robot model (1), and define $\hat{x} = x + Rcos\theta, \hat{y} = y + Rsin\theta$. Therefore,

$$\dot{\hat{x}} = vcos\theta - Rsin\theta\omega, \dot{\hat{y}} = vsin\theta - Rcos\theta\omega, \tag{9}$$

Now, we can linearize the robot model as (10), and the real control inputs are expressed as (11).

$$\dot{\hat{x}} = u_x, \dot{\hat{x}} = u_y, \tag{10}$$

$$\begin{bmatrix} v \\ \omega \end{bmatrix} = \begin{bmatrix} cos\theta & sin\theta \\ -\frac{sin\theta}{R} & \frac{cos\theta}{R} \end{bmatrix} + \begin{bmatrix} u_x \\ u_y \end{bmatrix},$$

$$\begin{bmatrix} v \\ \omega \end{bmatrix} = \begin{bmatrix} u_x cos\theta + u_y sin\theta \\ -\frac{u_x sin\theta}{R} + \frac{u_y cos\theta}{R} \end{bmatrix}, \tag{11}$$

The moving velocity for the ith robot during formation to a certain shape following certain predetermined track can be expressed as (12) based on (2),

$$q_i(t) = q_0(t) + \sum_{j=1}^{2} \alpha(s_{ij}(t_k^s), P_j(t_k^s)), \tag{12}$$

This can be considered as the derivative value based on the robot's relative velocity. For each robot in the system, we can assume the velocity of the robot to be a constant, and others follow the robot based on (12).

The ith robot's control design without time-delays is based on (2)(3)(4)(8)(12), and is given by,

$$u_i = \sum_{j=1}^{n} \alpha(s_{ij}(t_k^s), P_j(t_k^s)) sign(P_j(t) - P_i(t)) + \dot{q}_i(t), \tag{13}$$

in (13), $\alpha(s_{ij}(t_k^s)$ is the nonlinear control gain. $S(t_k^s)$ can be changing to reflect different types of communication strategies such as leader-follower, global communication and neighbor-follower, etc.

The design of nonlinear control gain for global communication and simple directed communication without time-delay can be given as,

$$\alpha(s_{ij}(t_k^s)) = \frac{s_{ij}(t_k^s)}{\sum_{l=1} s_{il}(t_k^s)}, \tag{14}$$

which has been proved in [1]. In this paper, we show through computer simulations and robots experiments that, the control gain (14) is sufficient for systems with directed communication topologies. However, this control gain may fail to achieve consensus when applied to systems with undirected communication topologies. Therefore, we adopt the new nonlinear control law for undirected communication as follows [1]. Let the nonlinear control gain α_{ij} be designed as,

case 1: if $P_i(t_k^s) = max_{j \in N_i} P_j(t_k^s) = min_{j \in N_i} P_j(t_k^s), \alpha_{ij}$can be any bounded positive value.

case 2: if $P_i(t_k^s) \geq max_{j \in N_i} P_j(t_k^s)$, then $alpha(s_{ij}$ can be ranged,

$$0 \leq \sum_{j \in N_i}^{n} \alpha(s_{ij}(t_k^s) < \frac{P_i(t_k^s) - min_{j \in N_i} P_j(t_k^s)}{c}, \tag{15}$$

case 3: if $P_i(t_k^s) \leq min_{j \in N_i} P_j(t_k^s)$, then $alpha(s_{ij}$ can be selected,

$$0 \leq \sum_{j \in N_i}^{n} \alpha(s_{ij}(t_k^s) < \frac{max_{j \in N_i} P_j(t_k^s) - P_i(t_k^s)}{c}, \tag{16}$$

case 4: if $min_{j \in N_i} P_j(t_k^s) < P_j(t_k^s) < max_{j \in N_i} P_j(t_k^s)$, let $alpha(s_{ij}$ be selected to satisfy,

$$0 \leq \sum_{j \in N_i}^{n} \alpha(s_{ij}(t_k^s) < min[(9), (10)], \tag{17}$$

where c could be any positive constant. This theorem has been proved in [1] analytically.

IV. Simulation and Experiments

A. Software Results

In the Matlab environment, we demonstrate the new control algorithm for multiagent systems. All the demos consist of four agents performing the consensus and formation control tasks with two communication strategies. One is the directed communication topology, which means each agent knows the connected neighbor agents' information in terms of velocity, position and orientation. The other is called least restrictive communication topology, in which each agent knows their neighbors' information in one way only.

To implement the control law shown in (13), for the directed communication topology, we consider a group of four mobile robots achieving a rhombus formation while following a circular movement. For the general setting as expressed in (12) (13), we set $q_0(t) = [2cost, 2sint]^T$, and $e_1(t) = [-sint, cost]^T, e_2(t) = [-cost, -sint]^T$. The four robots are following a circle, and the control gain applied is (14). The update time $t_s = 0.05s$. First, the control input (13) is adopted, which contains the sign function. Second, to simulate the control law without the sign function, simply remove the sign function part.

In the first set of simulations, the directed communication topology is adopted and we compare the system performance with and without the sign function. Four robots are designed to either converge to one point or form a certain shape (while making circular movements). $s_{ij}(t_k)$ is applied as,

$$s1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 \end{bmatrix}, s2 = \begin{bmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$s3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, s2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{18}$$

The obtained trajectories are shown in figures 1-4,

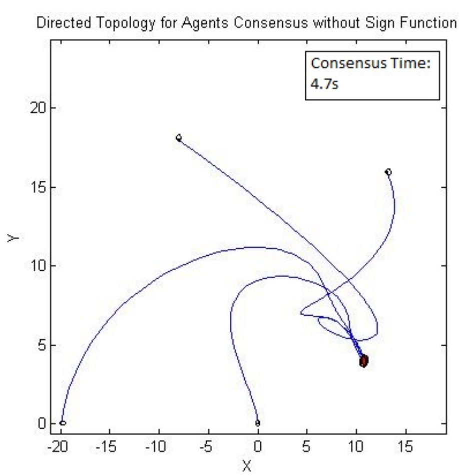

Fig. 1: Consensus of four agents without the sign function with directed communication topology, starting from four different positions and converging to one point. Time to consensus is T=4.7s

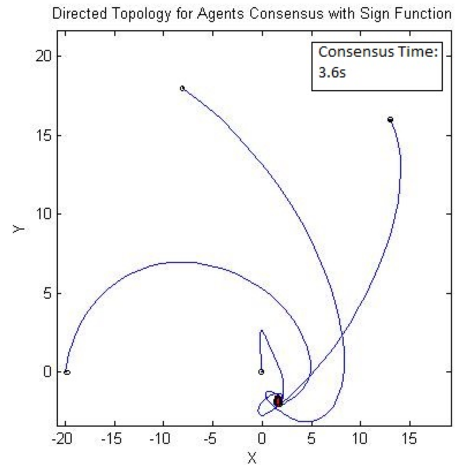

Fig. 2: Consensus of four agents with the sign function with directed communication topology, starting from four different positions and converging to one point. Time to consensus is T=3.6s

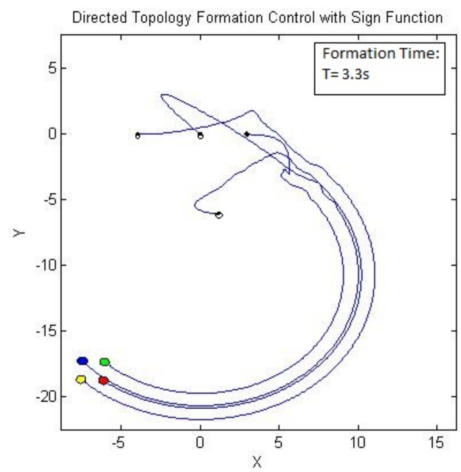

Fig. 3: Four robots starting from different positions and forming a rhombus while following a circle with the sign function and the directed communication topology. Time to the desired formation is T=3.3s

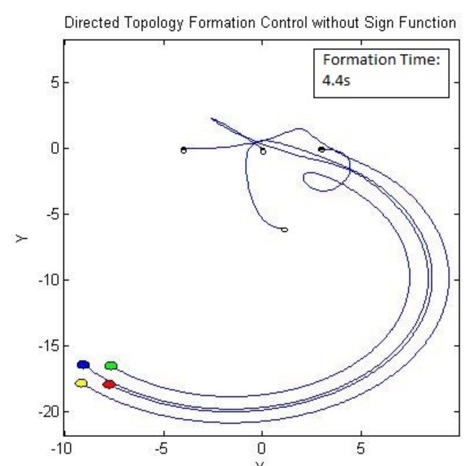

Fig. 4: Four robots starting from different positions forming a rhombus while following a circle without the sign function and with the directed communication topology. Time to the desired formation is t=4.4s

For the undirected communication topology, we adopt the least-restrictive neighbor communication topology, with each robot only communicating with a neighboring agent. There is no leader in the system, and minimum data transferring is incurred. $s_{ij}(t_k)$ is set as,

$$s1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}, s2 = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$s3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, s2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{19}$$

The control law with time-delays and the sign function can be expressed as (21), with τ_{ij} set to one second,

$$u_i = \sum_{l=1}^{n} \alpha(s_{ij}(t_k^s), P_j(t_k^s) - \tau_{ij}) sign(P_j(t - \tau_{ij}) - P_i(t)) + \dot{q}_i(t), \tag{21}$$

The control gain design is specified in (14)(15)(16)(17), i.e., every time an agent receives a neighbor's position information, the software compares this value with the maximum or minimum value as shown in (14)(15)(16)(17), then the control gain is selected accordingly.

As we mentioned before, the control gain in (14) may not be a appropriate for undirected communication topologies. Agents may fail to form the desired shape. Figure 5 below gives an example of such scenario with the continuous control gain (14). The four agents fail to converge within 60s and the trajectory contains oscillation.

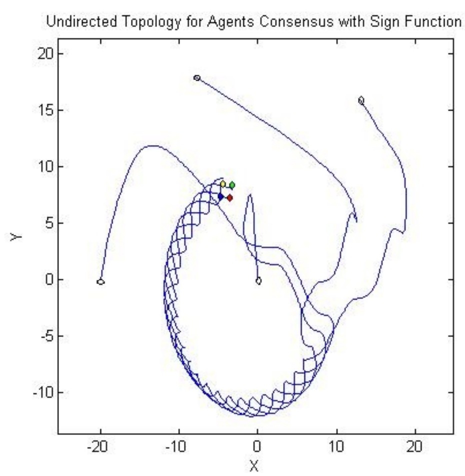

Fig. 5: Consensus of four agents with the sign function and time-delays with the directed communication topology. They start from four different positions and attempt to converge to one point, but fail to converge within 60s.

As we apply the rules (14)(15)(16)(17), the system works satisfactorily under the same conditions, as shown in figure 6. A formation example is also displayed in figure 7,

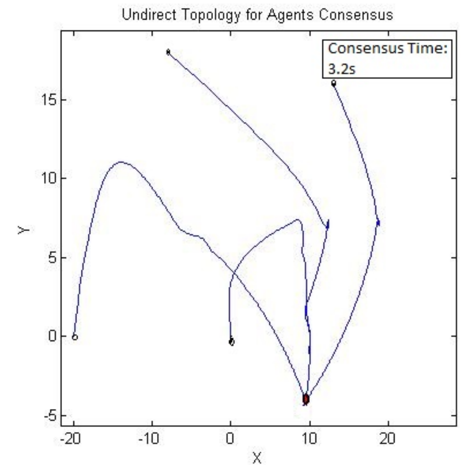

Fig. 6: Consensus of four agents using new rules with the sign function and time-delays with the directed communication topology. They start from four different positions to converge to one point in 3.2s.

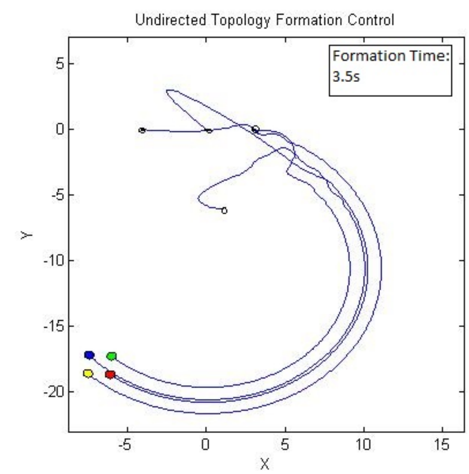

Fig. 7: Four robots starting from random positions form a rectangle following a certain circular movement. Time to the desired formation t=3.5s

B. Hardware Experiments

We also implemented the new control law in the Aira mobile robots. The robots know the initial position of themselves, but when they are moving, each robot only knows the velocity and the position information from one of its neighboring robots. $\dot{q}_i(t)$ is set based on the requirement of the consensus speed. All the Aira demos shown are for the distributed formation control of multiagent systems with switching topologies. $\dot{q}_i(t)$ is set to $200mm$ in both x and y orientations. The communication topology (19) is adopted with the control law (13). At this time, the control gains are designed as Kx=120 and Ky=20. Figure 8 shows the position information of robot1 and figure 9 compares the experimental results with the theory.

From figure 9 we can see the experimental value Kx=120 fits the theoretical control gain very well based on (15). Similarly, for the y domain, the experimental value fits well with the theoretical value (17).

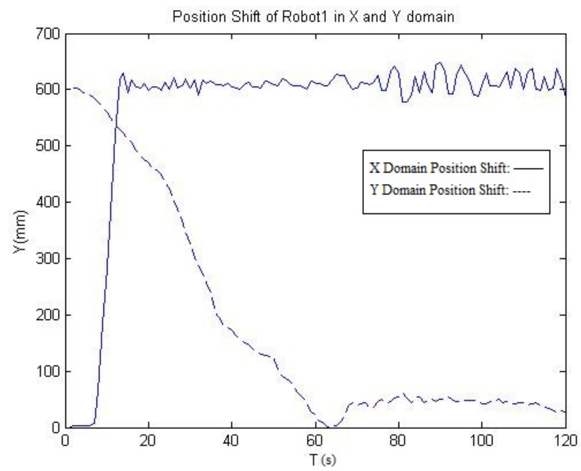

Fig. 8: Referring to Robot1 and Robot2, Y is the distance between two robots in x and y domains. We sampled 120 points, one point per second.

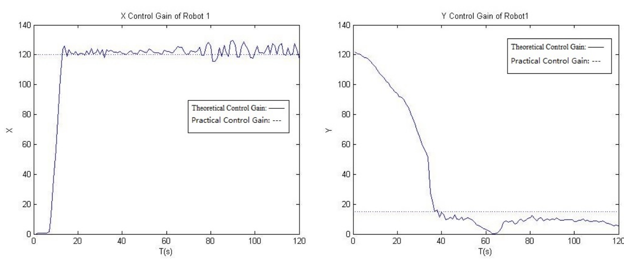

Fig. 9: X and Y control gains comparison between theory and experiments. We sampled 120 points, one point per second.

Figures 10 and 11 are a series of images showing the Aira robot simulations in our lab.

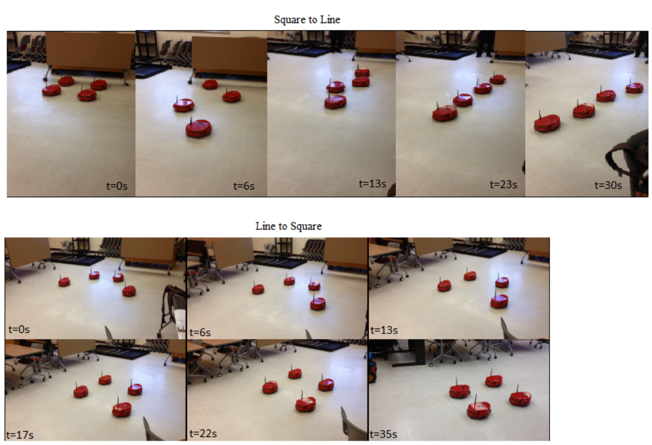

Fig. 10: Rectangle-to-line and line-to-rectangle formation control with undirected communication and the sign function.

Figure 11 shows the formation shape changing from rectangle to line, then to rhombus, and finally converging to one point. K is set to 60, moving velocity $\dot{q}_i(t) = 100mm$ with the same communication topology (19).

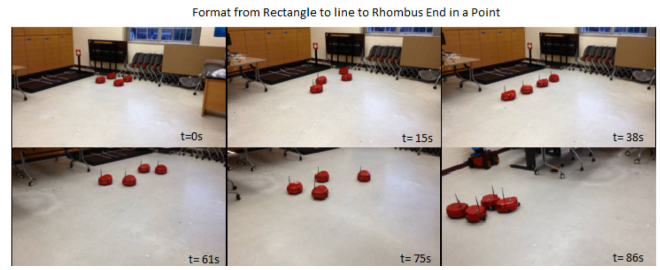

Fig. 11: Formation changes from rectangle to line to rhombus and ends with a point with undirected communication and the sign function

V. Conclusion

In this paper, we demonstrate the distributed cooperative control for the consensus of multigent systems with switching topologies and time-delays using Matlab and Aira mobile robot experiments. Matlab simulations confirmed the advantages of using the sign function in terms of short consensus time and assurance of consensus, compared with the control law without the sign function. Also, the results obtained through the discontinuous design of control gains validated the effectiveness of such design in consensus and formation problems. Another important benefit is, this new control law can be easily loaded into real robots, so practical implementation issues can be studied experimentally, such as time-delays, hardware limitations, etc.

References

[1] J. Wang, M. Obeng, Z. Qu, T. Yang, G. Staskevich and B. Abbe, Discontinuous cooperative control for consensus of multiagent systems with switching topologies and time-delays. Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on, , 0743-1546, 2013.

[2] J.Wang, M. Obeng, Distributed formation control for nonholonomic mobile robots. Sensor Array and Multichannel Signal Processing Workshop (SAM), 2012 IEEE 7th, 1551-2282, 2012.

[3] Z. Lin, B. Francis, and M. Maggiore, State agreement for continuous time coupled nonlinear systmes, SIAM Journal on Control and Optimization, vlo. 49, pp. 288-307, 2007.

[4] J. Wang, Z. Qu, C.M. Ihlefeld, and R.A. Hull, A control design based solution to robotic ecology: Autonomy of achieving cooperative behavior from a high-level astronaut command, Autonomous Robots, vol. 20, pp. 97-112, 2006.

[5] R. O. Saber, J. A. Fax, and R. M. Murray, Consensus and cooperation in networked multiagent systems, Proceedings of the IEEE, vol. 95, pp. 215-233, 2007.

[6] Q. Hui, Finite time rendezvous algorithms for mobile autonomous agents, IEEE Trans. Automa. Control , vol. 56, pp. 1213-1223, 2008.

[7] B. Paden and S. Sastry, A calculus for computing filippov's differential inclusion with application to the variable structure control of robot mainpulators, IEEE Transactions on Circuits and systems, vol. 34, pp. 73-82, 1987.

[8] N. Moshtagh and A. Jadbabaie. Distributed geodesic control laws for flocking of nonholonomic agents. IEEE Transactions on Automatic Control, 52:681-686, 2007.

[9] Z. Qu, Cooperative control of dynamical systems. Springer-Verlag, London, 2009.

[10] M. Arcak. Passivity as a design tool for group coordination. IEEE Trans. Automat. Contr., 52:1380-1389, 2007.

[11] J. Wang, Z. Qu, and M. Obeng. A distributed cooperative steering control with application to nonholonomic robots. In 49th IEEE conference on Dec. and Ctrl, pages 4571-4576, Atlanta, GA, Dec 2010.

[12] W. Ren and R. W. Beard. Distributed consensus in multi-vehicle cooperative control. Springer-Verlag, London, 2008.

[13] A. Jadbabaie, J. Lin, and A.S. Morse. Coordination of groups of vehicle formations. IEEE transactions on automatic control. 49: 1465-1476, 2004

[14] F. Bullo, J. Cortes, and S. Martinez, Distributed control robotic networks, ser. Applied Mathematics Series. Princeton University Press, 2009, electronically available at http://coordinationbllk.info.

[15] Z. Butler and D. Rus, ”Controlling mobile sensors for monitoring events with coverage constrains,” in Proceedings of IEEE International Conference of Robotics and Automation, New Orleans, LA, April 2004, pp. 1563-1573

[16] M. Schwager, J. McLukrin, and D. Rus, ”Distributed coverage control with sensory feedback for networked robots,” in Proceedings of Robotics: Science and Systems, Philadelphia, PA, August 2006.

[17] D.V. Dimarogonas and K. J. Kyriakopoulos, ”Connectedness preserving distributed swarm aggregation for multiple kinematic robots.” IEEE Transactions on Robotics, vol. 24, pp. 1213-1223, 2008.

[18] Z. Lin, M. Brouchke, and B. Francis, ”Local control strategies for groups of mobile autonomous agents using nearest neighbor rules,” IEEE Trans. on Automatic Control, vol.622-629, 2004.

[19] W. Wang and J.J.E. Slotine, ”A theoretical study of different leader roles in networks,” IEEE Transactions on Automatic Control, vol. 51, no.7, pp. 1156-1161, July 2006.

[20] G. T. Sibley, M. H. Rahimi, and G. S. Sukhatme, ”Robomote: A tiny mobile robot platform for large-scale sensor networks,” in Proceedings of IEEE International Conference on Robotics and Automation(ICRA), 2002.

[21] R. Olfati-Saber and R. R. Murray,”Consensus problems in networks of agents with switching topology and time-delays,” IEEE Transactions on Automatic Control, vol. 48, no.9, pp. 1520-1533, September 2004.

Formation Control of Multiple Nonholonomic Mobile Robots with Limited Information of a Desired Trajectory

Jing Wang, Morrison Obeng, Tianyu Yang, Gennady Staskevich, and Brian Abbe

Abstract— **In the study of task coordination for multiagent systems, formation control has received considerable attention due to its potential applications in civil and/or military practices. Fundamentally, formation control problem for multiagent systems can be formulated as making a group of agents follow the desired trajectory while maintaining certain prescribed geometric distances among agents. In this paper, we consider the formation control problem for mobile robots with nonlinear dynamics and moving in a 2D environment. To address the inherent challenges due to nonlinear system dynamics and agents' limited sensing/communication capabilities, we instill an idea of integrating the recently developed distributed consensus theory into the standard feedback control, and propose a new time-varying cooperative control strategy to solve the formation control problem for multiagent systems. In particular, the proposed design only requires the local and intermittent information exchange among agents to achieve the formation control objective. More importantly, we remove the restriction on the need of the desired trajectory for every agent, and instead design a distributed observer for obtaining the desired trajectory in order to establish the formation in the design. The overall distributed formation control system stability is rigorously proved by using a contraction mapping method under the condition that the sensing/communication network among robots is sequentially complete. Simulation is provided to validate the effectiveness of the proposed design.**

I. Introduction

Coordination of multiagent systems has been and continues to be an active research area in the current control community, and recent years have seen a significant progress in the development of distributed consensus strategies for multiagent systems [24][23][25].

The developments have been primarily focused on addressing two aspects of issues. The first issue is on the study of network controllability. The core is to identify the least required sensing/communication conditions for completing coordination tasks. A significant result was obtained in [10], in which the sensing/communication topologies are modeled using an undirected graph and its uniform connectivity provides the sufficient condition for agents consensus. This condition was further relaxed to take into account the directed graph [25][27], and the existence of a spanning tree in the graph is necessary and sufficient for group coordination. In our recent work [24], we addressed this issue by using matrix theory, and introduced the notion of sequential completeness of sensing/communication matrix sequences to describe the connectivity condition. The second issue is on the design of cooperative control and stability analysis. The cooperative control for linear systems are thoroughly studied in [13][10][27][8][24]. For nonlinear systems, some results are available by using passivity-based design in [2], Lyapunov design in [23], and set-valued Lyapunov functions [19][16].

The aforementioned results are mostly centered on the study of consensus problem for multiple dynamical systems. One of direct applications of those results is to tackle the formation control problem for real systems. Fundamentally, formation control problem can be formulated as making a group of agents follow the desired trajectory while maintaining certain prescribed geometric distances among agents. In this paper, we consider the formation control problem for nonholonomic mobile robots with nonlinear dynamics and moving in a 2D environment. Progress has been made in solving formation control problem by using leader-follow strategies, virtual structure method, and artificial potentials [26][11][22][7][5][17][9][1][14][18]. However, most results are obtained by either assuming linear system dynamics or converting the nonholonomic robot model into linear one through feedback linearization. It is well known that it is challenging to design feedback control for mobile robots with nonholonomic constraints [21]. There are some results for studying the cooperative control of nonholonomic robots [6][15][20][28]. Among them, a discontinuous control was proposed in [6] and nonsmooth Lyapunov theory and graph theory are used for stability analysis. In [15], based on the Frenet-Serret model of unicycle, time-varying controls were designed and analyzed using average theory. The work in [20][28] assumed the constant driving velocity and controls were only designed for steering velocity. In our recent paper [30], we proposed a distributed cooperative steering control design for a class of networked dynamical systems with inherent nonlinear dynamics. A number of conditions were established in terms of the properties of the cooperative steering control for achieving cooperative behaviors.

In this paper, we present a new solution for the formation control design of nonholonomic mobile robots. The nonholonomic constraints of robots are explicitly taken into account in the proposed design by converting the unicycle model into the canonical chained form. The proposed distributed formation control utilizes the information exchange to better coordinate the motion of individual robots, but there is no requirement for the strongly connected communication topology which could be uncertain and unreliable due to

J. Wang and M. Obeng are with Faculty of Computer Engineering, School of Science, Engineering and Mathematics, Bethune-Cookman University, Daytona Beach, FL 32114, USA. G. Staskevich and B. Abbe are with AFRL/RISC, Rome, NY, 13441. T. Yang is with Electrical Engineering, Embry-Riddel Aeronautical University, Daytona Beach, FL 32114. Email: ecejwang@ieee.org This work was supported by the Air Force Research Laboratory FA8750-13-1-0109.

communication noises. Instead, we allow more flexible, intermittent, and time varying communication topologies among robots. A finite-time distributed observer is designed to estimate the desired trajectory information. The stability and convergence analysis for the proposed formation control are done through contraction mapping method under the condition that the sensing/communication network among robots is sequentially complete. Simulation result is included to validate the effectiveness of the proposed control.

The rest of the paper is organized as follows. Section II formulates the formation control problem. Main results are presented in section III, in which distributed finite-time observer and nonlinear formation control are designed including system stability analysis. A simulation example is given in section IV. Section V concludes the paper.

II. Problem Formulation

Consider a network of multiple nonholonomic mobile robots with the individual system dynamics given by

$$\begin{aligned} \dot{x}_i &= v_i \cos\theta_i, \\ \dot{y}_i &= v_i \sin\theta_i, \\ \dot{\theta}_i &= \omega_i \end{aligned} \tag{1}$$

where $i \in \Omega \triangleq \{1, \cdots, n\}$, $(x_i, y_i) \in \Re^2$ denotes the ith robot's position, θ_i is the orientation, $v_i \in \Re$ driving velocity, and $\omega_i \in \Re$ the steering velocity.

To make the method in this paper more general, our design will be based on the following canonical chained form

$$\dot{z}_{i1} = u_{i1}, \quad \dot{z}_{i2} = u_{i2}, \quad \dot{z}_{i3} = z_{i2}u_{i1}, \tag{2}$$

into which the model in (1) can be converted by using the following state and input transformations

$$z_{i1} = x_i, \quad z_{i2} = \tan\theta_i, \quad z_{i3} = y_i, \tag{3}$$

$$u_{i1} = v_i \cos\theta_i, \quad u_{i2} = \frac{\omega_i}{\cos^2\theta_i}. \tag{4}$$

That is, once the formation controls u_{i1} and u_{i2} are designed, the corresponding v_i and ω_i can be found through the inverse transformation of (4).

The design objective of this paper is to coordinate the motion of individual robots to follow a desired trajectory contour while maintaining certain prescribed geometric formation shape through local information exchange among robots. By taking the whole group of mobile robots as a virtual body moving along the desired trajectory, formation shape of robots in the group can be determined by a set of local coordinates with reference to the moving frame attached to the desired trajectory (see figure 1 for an illustration of moving frames).

More specifically, let $q_0(t) = [x_0(t), y_0(t)]^T \in \Re^2$ be the desired trajectory for the group motion, the moving frame $\mathcal{F}(t)$ attached to $q_0(t)$ can be defined by the following orthonormal vectors $e_1(t)$ and $e_2(t)$

$$e_1(t) = \begin{bmatrix} e_{11}(t) \\ e_{12}(t) \end{bmatrix} = \begin{bmatrix} \dfrac{\dot{x}_0(t)}{\sqrt{[\dot{x}_0(t)]^2 + [\dot{y}_0(t)]^2}} \\ \dfrac{\dot{y}_0(t)}{\sqrt{[\dot{x}_0(t)]^2 + [\dot{y}_0(t)]^2}} \end{bmatrix},$$

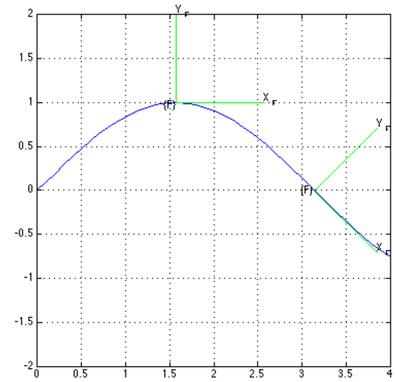

Fig. 1. Moving frames on the desired trajectory

$$e_2(t) = \begin{bmatrix} e_{21}(t) \\ e_{22}(t) \end{bmatrix} = \begin{bmatrix} -\dfrac{\dot{y}_0(t)}{\sqrt{[\dot{x}_0(t)]^2 + [\dot{y}_0(t)]^2}} \\ \dfrac{\dot{x}_0(t)}{\sqrt{[\dot{x}_0(t)]^2 + [\dot{y}_0(t)]^2}} \end{bmatrix}.$$

Accordingly, any formation consisting of n robot positions in $\mathcal{F}(t)$ can be expressed as $\{P_1, \cdots, P_n\}$ with

$$P_i(t) = \alpha_{i1}e_1(t) + \alpha_{i2}e_2(t), \tag{5}$$

where α_{ij} are constants of determining the formation. To this end, the formation control objective can be recast as to design the control laws $v_i(t)$ and $\omega_i(t)$ for the ith robot such that

$$\lim_{t\to\infty} \left[\begin{bmatrix} x_i(t) \\ y_i(t) \end{bmatrix} - q_0(t) - P_i(t) \right] = 0. \tag{6}$$

It is noted that the control objective defined in (6) can be achieved through the standard tracking control design for individual robots if the desired trajectory $q_0(t)$ and its derivative $\dot{q}_0(t)$ are available to every robot. However, such a design may not be robust in the presence of disturbance and noise measurements due to the lack of coordination among robots. On the other hand, the desired trajectory $q_0(t)$ may be known only by some of robots in the group. Therefore, it is desirable to design distributed formation control law for the ith robot based on information exchange and relative position measurement between robots within its sensing/communication range.

In this paper, we assume that the sensing/communication topologies among robots are changing, which can be captured by the time sequence $\{t_\eta^s : \eta = 0, 1, \cdots\}$. Correspondingly, we introduce the following binary matrix to describe the sensing/communication topology.

$$S(t) = \begin{bmatrix} s_{11} & s_{12}(t) & \cdots & s_{1q}(t) \\ s_{21}(t) & s_{22} & \cdots & s_{2q}(t) \\ \vdots & \vdots & \ddots & \vdots \\ s_{n1}(t) & s_{n2}(t) & \cdots & s_{nn} \end{bmatrix}, \tag{7}$$

with $S(t) = S(t_\eta^s), \forall t \in [t_\eta^s, t_{\eta+1}^s)$, where $s_{ii} \equiv 1$; $s_{ij}(t) = 1$ if the jth robot is in the sensing/communication range

of the ith robot at time t, and $s_{ij} = 0$ if otherwise; and $t_0^s \triangleq t_0$. It can be assumed without loss of any generality that $0 < \underline{c}_t \le t_{\eta+1}^s - t_\eta^s \le \overline{c}_t < \infty$, where $\underline{c}_t$ and $\overline{c}_t$ are constant bounds. We also define the neighbor set for the ith robot as $\mathcal{N}_i(t) = \{j \in \Omega | s_{ij}(t) = 1\}$.

It is apparent that in order to achieve the coordination task, the sensing/communication topologies defined by (7) must satisfy certain connectivity conditions. In our recent work [24][30], we introduced the notion of *sequentially completeness* to describe the least required condition on network connectivity for cooperative control design, which is restated by the following definitions.

Definition 2.1: Sensing/communication matrix sequence $\{S(t)\}$ is said to be *sequentially lower-triangularly complete* if it is sequentially lower-triangular and in every row i of its lower triangular canonical form, there is at lease one $j < i$ such that the corresponding block $S_{ij}(t)$ is uniformly non-vanishing.

Definition 2.2: Sensing/communication matrix sequence $\{S(t)\}$ is said to be *sequentially complete* if the sequence contains an infinite subsequence that is sequentially lower-triangularly complete.

As an example for sequential completeness, let us assume that the sensing/communication topologies for 3 robots are changing according to the sequence $\{S(t_k), k \in \aleph, \aleph \triangleq \{1, 2, \dots\}\}$ defined below: $S(t_k) = S_1$ for $k = 4\eta$, $S(t_k) = S_2$ for $k = 4\eta + 1$, $S(t_k) = S_3$ for $k = 4\eta + 2$, and $S(t_k) = S_4$ for $k = 4\eta + 3$, where $\eta \in \aleph$,

$$S_1 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad S_2 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$
$$S_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \text{ and } S_4 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{8}$$

The bitwise union of $S_i, i = 1, \cdots, 4$ is

$$\bigcup_i S_i = \left[\begin{array}{cc|c} 1 & 1 & 0 \\ 1 & 1 & 0 \\ \hline 1 & 0 & 1 \end{array}\right] \triangleq \left[\begin{array}{c|c} S'_{\Lambda,11} & \emptyset \\ \hline S'_{\Lambda,21} & 1 \end{array}\right].$$

It then follows from the structure of $\bigcup_i S_i$ that the corresponding sequence is lower-triangularly complete, and therefore the switching sensing/communication topologies defined by (8) is *sequentially complete*.

Assumption 2.1: The group of robots defined in (1) has a sequentially complete sensing/communication network.

III. The Main Result

In this section, we present a nonlinear formation control design for nonholonomic mobile robots (1) with limited information of the desired trajectory $q_0(t)$. Particularly, the proposed new formation control will be done with the aid of distributed observers for the estimation of $q_0(t)$.

It follows that the desired trajectory $q_0(t)$ also satisfies the nonholonomic constraints, that is, we have

$$\dot{x}_0 = v_0 \cos\theta_0, \quad \dot{y}_0 = v_0 \sin\theta_0, \quad \dot{\theta}_0 = \omega_0$$

for some v_0, ω_0, and θ_0. It is then readily seen that the moving frame attached to $q_0(t)$ can be established using the rotation matrix in terms of the desired trajectory orientation $\theta_0(t)$, that is,

$$e_1(t) = \begin{bmatrix} \cos\theta_0 \\ \sin\theta_0 \end{bmatrix}, e_2(t) = \begin{bmatrix} -\sin\theta_0 \\ \cos\theta_0 \end{bmatrix}$$

To this end, the formation control can be designed based on the real time estimation of $x_0(t), y_0(t)$ and $\theta_0(t)$. The proposed distributed observer is of the form (for $t \in [t_k^s, t_{k+1}^s)$)

$$\begin{aligned} \dot{x}_{i,0}(t) &= \sum_{j\in\mathcal{N}_i} \alpha_{ij} \mathrm{sgn}(x_{j,0}(t_k^s) - x_{i,0}(t_k^s)) \\ &\quad + \alpha_{i0} s_{i0} \mathrm{sgn}(x_0(t_k^s) - x_{i,0}(t_k^s)) \quad (9) \\ \dot{y}_{i,0}(t) &= \sum_{j\in\mathcal{N}_i} \alpha_{ij} \mathrm{sgn}(y_{j,0}(t_k^s) - y_{i,0}(t_k^s)) \\ &\quad + \alpha_{i0} s_{i0} \mathrm{sgn}(y_0(t_k^s) - y_{i,0}(t_k^s)) \quad (10) \\ \dot{\theta}_{i,0}(t) &= \sum_{j\in\mathcal{N}_i} \alpha_{ij} \mathrm{sgn}(\theta_{j,0}(t_k^s) - \theta_{i,0}(t_k^s)) \\ &\quad + \alpha_{i0} s_{i0} \mathrm{sgn}(\theta_0(t_k^s) - \theta_{i,0}(t_k^s)) \quad (11) \end{aligned}$$

where $x_{i,0}(t)$, $y_{i,0}(t)$ and $\theta_{i,0}(t)$ are the ith robot's estimate of $x_0(t)$, $y_0(t)$, and $\theta_0(t)$, respectively, $s_{i0} = 1$ if and only if the ith robot has the direct access to the information of the desired trajectory, $\alpha_{i,j}$ and α_{i0} are piecewise constant control gains to be designed, and sgn$(\cdot)$ function is defined as

$$\mathrm{sgn}(z) = \begin{cases} 1, & z > 0 \\ 0, & z = 0 \\ -1, & z < 0 \end{cases}$$

The following theorem states the finite-time convergence of the proposed distributed observers (9), (10) and (11) under appropriate choices of $\alpha_{i,j}$ and α_{i0}.

Theorem 1: Consider a group of nonholonomic mobile robots given by (2) with assumption 2.1. The finite-time convergence of $x_{i,0}(t)$ to $x_0(t)$, $y_{i,0}(t)$ to $y_0(t)$ and $\theta_{i,0}(t)$ to $\theta_0(t)$ can be guaranteed under the proposed distributed observers (9), (10) and (11), if the control gain α_{ij} and α_{i0} (for $s_{i0} = 1$) are designed as follows: for agent i,

$$\sum_{j\in\mathcal{N}_i} a_{ij}(t_k^s) + a_{i0}(t_k^s) > \bar{d}, \text{ if } s_{i0} = 1, \tag{12}$$
$$\sum_{j\in\mathcal{N}_i} a_{ij}(t_k^s) > \bar{d}, \text{ if } s_{i0} = 0 \tag{13}$$

where $\bar{d}$ is the upper bound of $|\dot{x}_0(t)|, |\dot{y}_0(t)|$ and $|\dot{\theta}_0(t)|$.

Proof: We prove the convergence of (9). The same procedure applies to (10) and (11). We first consider the robots which have the direct access to $x_0(t)$, that is, robot i for $i \in \Omega_0 = \{i \in \Omega : s_{i0} = 1\}$. Define $x_{i,0}^{max}(t) = \max_{i\in\Omega_0} x_{i,0}(t)$ and $x_{i,0}^{min}(t) = \max_{i\in\Omega_0} x_{i,0}(t)$. Let $\tilde{x}_{i,0}^{max}(t) = x_{i,0}^{max}(t) - x_0(t)$, and $\tilde{x}_{i,0}^{min}(t) = x_{i,0}^{min}(t) - x_0(t)$. It follows from (9)

that

$$\dot{\tilde{x}}_{i,0}^{max}(t) = \sum_{j\in\mathcal{N}_i}\alpha_{ij}\text{sgn}(\tilde{x}_{j,0}(t_k^s)-\tilde{x}_{i,0}^{max}(t_k^s)) +\alpha_{i0}\text{sgn}(x_0(t_k^s)-x_{i,0}^{max}(t_k^s))-\dot{x}_0 \quad (14)$$

$$\dot{\tilde{x}}_{i,0}^{min}(t) = \sum_{j\in\mathcal{N}_i}\alpha_{ij}\text{sgn}(\tilde{x}_{j,0}(t_k^s)-\tilde{x}_{i,0}^{min}(t_k^s)) +\alpha_{i0}\text{sgn}(x_0(t_k^s)-x_{i,0}^{min}(t_k^s))-\dot{x}_0 \quad (15)$$

Now consider the evolution of $\tilde{x}_{i,0}^{max}(t)$ and $\tilde{x}_{i,0}^{min}(t)$. If both $\tilde{x}_{i,0}^{max}(t) > 0$ and $\tilde{x}_{i,0}^{min}(t) \geq 0$, it suffices to show that $\tilde{x}_{i,0}^{max}(t)$ converges to zero in finite time. It follows from (14) and (12) that

$$\begin{aligned}\dot{\tilde{x}}_{i,0}^{max}(t) &= -\sum_{j\in\mathcal{N}_i}\alpha_{ij}-\alpha_{i0}-\dot{x}_0\\ &\leq -\sum_{j\in\mathcal{N}_i}\alpha_{ij}-\alpha_{i0}+\bar{d}<0, \quad (16)\end{aligned}$$

which implies that $\tilde{x}_{i,0}^{max}$ will converge to zero in a finite time. If both $\tilde{x}_{i,0}^{max}(t) \leq 0$ and $\tilde{x}_{i,0}^{min}(t) < 0$, it suffices to show that $\tilde{x}_{i,0}^{min}(t)$ converges to zero in finite time. It follows from (15) and (12) that

$$\begin{aligned}\dot{\tilde{x}}_{i,0}^{min}(t) &= \sum_{j\in\mathcal{N}_i}\alpha_{ij}+\alpha_{i0}-\dot{x}_0\\ &\geq \sum_{j\in\mathcal{N}_i}\alpha_{ij}+\alpha_{i0}-\bar{d}>0, \quad (17)\end{aligned}$$

which implies that $\tilde{x}_{i,0}^{min}$ will converge to zero in a finite time. If $\tilde{x}_{i,0}^{max}(t) > 0$ and $\tilde{x}_{i,0}^{min}(t) < 0$, same conclusion can be drawn by noting both (16) and (17).

Now let us consider the robots which do not have direct access to $x_0(t)$ but have information exchange with robots in Ω_0, that is, for robot i with $i \in \Omega_1 = \{i \in \Omega : s_{i0} = 0, s_{ij} = 1, j \in \Omega_0\}$. To this end, we know that since after certain finite time, robot j (for $j \in \Omega_0$) will converge to $x_0(t)$, thus, similar analysis can be done to show the convergence of $\tilde{x}_{i,0}$ (for $i \in \Omega_1$) to zero in finite time. The above procedure can be repeated recursively due to the sequentially complete sensing/communication assumption. □

Remark 3.1: The proposed distributed observers are motivated by the result in [4], in which a distributed estimator with terms like $\text{sgn}(\sum_{j=1}^n a_{lj}(x_j - x_l))$ was proposed, while in this paper we consider more general form with terms like $\text{sgn}(x_j - x_l)$. ◇

In what follows, we present the distributed formation control design. Let us define an infinite sequence of time instants $\{t_0 + kT_s\}$ for $k \in \mathcal{N} \triangleq \{0, 1, \cdots\}$ and with sampling time $0 < T_s \leq \underline{c}_t$. The control inputs will be updated according to the sampling time instants. For notational convenience, $z(t_0 + kT_s)$ is simplified as $z(k)$ for any variable z.

Theorem 2: Consider a group of nonholonomic mobile robots given by (2) with assumption 2.1. Let the distributed cooperative control be for $t \in [t_0 + kT_s, t_0 + (k+1)T_s)$

$$u_{i1}(t) = a_{i1}^k + a_{i2}^k \sin\omega(t-t_0-kT_s) \quad (18)$$

$$u_{i2}(t) = b_{i1}^k + b_{i2}^k \cos\omega(t-t_0-kT_s) \quad (19)$$

where $\omega = \frac{2\pi}{T_s}$, $a_{i2}^k \neq 0$ can be any constant, and

$$a_{i1}^k = \frac{1}{T_s}\sum_{j=1}^n G_{ij}(k)[x_j(k)-x_i(k)-x_j^d(k) +x_i^d(k+1)], \quad (20)$$

$$b_{i1}^k = \frac{1}{T_s}\sum_{j=1}^n G_{ij}(k)[z_{j2}(k)-z_{i2}(k)], \quad (21)$$

$$b_{i2}^k = \frac{2\omega}{a_{i2}^k T_s}\left[\sum_{j=1}^n G_{ij}(k)[y_j(k)-y_i(k)-y_j^d(k) +y_i^d(k+1)] - \frac{a_{i1}^k b_{i1}^k T_s^2}{2} -a_{i1}^k z_{i2}(k)T_s + \frac{a_{i2}^k b_{i1}^k T_s}{\omega}\right]. \quad (22)$$

with

$$G_{ij}(k) = \frac{s_{ij}(k)}{\sum_{\eta=1}^n s_{i\eta}(k)}, \quad j = 1, \cdots, n. \quad (23)$$

$$x_i^d(k) = x_{i,0}(k) + \alpha_{i1}\cos\theta_{i,0}(k) - \alpha_{i2}\sin\theta_{i,0}(k)$$

$$y_i^d(k) = y_{i,0}(k) + \alpha_{i1}\sin\theta_{i,0}(k) + \alpha_{i2}\cos\theta_{i,0}(k)$$

$$x_i^d(k+1) = x_{i,0}(k+1) + \alpha_{i1}\cos\theta_{i,0}(k+1) -\alpha_{i2}\sin\theta_{i,0}(k+1)$$

$$y_i^d(k+1) = y_{i,0}(k+1) + \alpha_{i1}\sin\theta_{i,0}(k+1) +\alpha_{i2}\cos\theta_{i,0}(k+1)$$

where $x_{i,0}$, $y_{i,0}$, and $\theta_{i,0}$ are governed by (9), (10), and (11), respectively, and

$$x_{i,0}(k+1) = x_{i,0}(k) +T_s(\sum_{j\in\mathcal{N}_i}\alpha_{ij}\text{sgn}(x_{j,0}(k)-x_{i,0}(k)) +\alpha_{i0}s_{i0}\text{sgn}(x_0(k)-x_{i,0}(k))) \quad (24)$$

$$y_{i,0}(k+1) = y_{i,0}(k) +T_s(\sum_{j\in\mathcal{N}_i}\alpha_{ij}\text{sgn}(y_{j,0}(k)-y_{i,0}(k)) +\alpha_{i0}s_{i0}\text{sgn}(y_0(k)-y_{i,0}(k))) \quad (25)$$

$$\theta_{i,0}(k+1) = \theta_{i,0}(k) +T_s(\sum_{j\in\mathcal{N}_i}\alpha_{ij}\text{sgn}(\theta_{j,0}(k)-\theta_{i,0}(k)) +\alpha_{i0}s_{i0}\text{sgn}(\theta_0(k)-\theta_{i,0}(k))) \quad (26)$$

Then the formation control objective (6) is achieved.

Proof: Directly applying controls (18) and (19) to (2) yields

$$z_{i1}(k+1) = z_{i1}(k) + a_{i1}^k T_s \quad (27)$$

$$z_{i2}(k+1) = z_{i2}(k) + b_{i1}^k T_s \quad (28)$$

$$z_{i3}(k+1) = z_{i3}(k) + a_{i1}^k z_{i2}(k)T_s + \frac{a_{i1}^k b_{i1}^k T_s^2}{2} -\frac{a_{i2}^k b_{i1}^k T_s}{\omega} + \frac{b_{i2}^k a_{i2}^k T_s}{2\omega} \quad (29)$$

Substituting (20), (21) and (22) into the above equations, and noting that a_{i1}^k and b_{i2}^k can be rewritten as

$$a_{i1}^k = \frac{1}{T_s}\sum_{j=1}^n G_{ij}(k)[x_{j1}(k) - x_{i1}(k) - x_j^d(k) + x_i^d(k)] + \frac{1}{T_s}(x_i^d(k+1) - x_i^d(k)), \quad (30)$$

$$b_{i2}^k = \frac{2\omega}{a_{i2}^k T_s}\left[\sum_{j=1}^n G_{ij}(k)[y_{j3}(k) - y_{i3}(k) - y_j^d(k) + y_i^d(k)] + (y_i^d(k+1) - y_i^d(k)) - \frac{a_{i1}^k b_{i1}^k T_s^2}{2} - a_{i1}^k z_{i2}(k) T_s + \frac{a_{i2}^k b_{i1}^k T_s}{\omega}\right], \quad (31)$$

we have

$$z_{i1}(k+1) = z_{i1}(k) + \sum_{j=1}^n G_{ij}(k)[\zeta_{j1}(k) - \zeta_{i1}(k)] + \int_{t_0+kT_s}^{t_0+(k+1)T_s} \dot{x}_i^d(t)dt, \quad (32)$$

$$z_{i2}(k+1) = z_{i2}(k) + \sum_{j=1}^n G_{ij}(k)[z_{j2}(k) - z_{i2}(k)], \quad (33)$$

$$z_{i3}(k+1) = z_{i3}(k) + \sum_{j=1}^n G_{ij}(k)[\zeta_{j3}(k) - \zeta_{i3}(k)] + \int_{t_0+kT_s}^{t_0+(k+1)T_s} \dot{y}_i^d(t)dt, \quad (34)$$

where $\zeta_{l1}(k) \triangleq x_l(k) - x_l^d(k)$ and $\zeta_{l3}(k) \triangleq y_l(k) - y_l^d(k)$. It then follows from the definitions of $\zeta_{l1}(k)$ and $\zeta_{l3}(k)$ and from (32) and (34) that

$$\zeta_{i1}(k+1) = \zeta_{i1}(k) + \sum_{j=1}^n G_{ij}(k)[\zeta_{j1}(k) - \zeta_{i1}(k)], \quad (35)$$

$$\zeta_{i3}(k+1) = \zeta_{i3}(k) + \sum_{j=1}^n G_{ij}(k)[\zeta_{j3}(k) - \zeta_{i3}(k)]. \quad (36)$$

To this end, if we can show that $\lim_{k\to\infty}\zeta_{i1}(k) = c_1$, $\lim_{k\to\infty}\zeta_{i3}(k) = c_3$, $\lim_{k\to\infty} z_{i2}(k) = c_2$, for all i, where c_1, c_2, c_3 are some constants, then it is obvious that the formation control objective (6) is achieved since $\lim_{k\to\infty} z_{i1}(k) = c_1 + x_i^d(k)$ and $\lim_{k\to\infty} z_{i3}(k) = c_3 + y_i^d(k)$.

It suffices to show the convergence of $\zeta_{i1}(k)$ in (35). Same argument can be applied to $\zeta_{i3}(k)$ in (36) and $z_{i2}(k)$ in (33). It follows from the choice of $G_{ij}(k)$ in (23) that

$$\sum_{j=1}^n G_{ij}(k)[\zeta_{j1}(k) - \zeta_{i1}(k)] = \sum_{j=1}^n G_{ij}(k)\zeta_{j1}(k) - \zeta_{i1}(k).$$

Thus, (35) becomes

$$\zeta_{i1}(k+1) = \sum_{j=1}^n G_{ij}(k)\zeta_{j1}(k) \quad (37)$$

To show the convergence of $\zeta_{i1}(k)$, let us define $\Omega = \{1, \cdots, n\}$ to be the set of indices on state variables, and at time instant $t_0 + kT_s$, let $\zeta_{\max}^1(k) = \max_j \zeta_{j1}(k)$ and $\zeta_{\min}^1(k) = \min_j \zeta_{j1}(k)$. Define sub-sets $\Omega_{\max}(k)$, $\Omega_{\min}(k)$, $\Omega_{\max}^c(k) = \Omega/\Omega_{\max}(k)$, and $\Omega_{\min}^c(k) = \Omega/\Omega_{\min}(k)$ as follows:

$$\Omega_{\max}(k) = \{i^* \in \Omega : \zeta_{i^*1}(k) = \zeta_{\max}^1(k)\}$$

and

$$\Omega_{\min}(k) = \{i_* \in \Omega : \zeta_{i_*1}(k) = \zeta_{\min}^1(k)\}$$

Thus, the proof of the convergence $\zeta_{i1}(k)$ is equivalent to prove for any k, there exists a constant $\delta(k) > 0$, such that

$$\|\zeta_{\max}^1(k+\delta) - \zeta_{\min}^1(k+\delta)\| \le \lambda\|\zeta_{\max}^1(t) - \zeta_{\min}^1(k)\|, \quad (38)$$

for some $0 \le \lambda < 1$.

Noting that It follows from the sampling time $T_s \le \underline{c}_t$ that the sensing/communication sequence $\{S(k) \triangleq S(t_0 + kT_s), k = 0, 1, \cdots\}$ completely captures the information of sequence $\{S(t_\eta^s), \eta = 0, 1, \cdots\}$. Thus, the sequential completeness of $\{S(t_\eta^s), \eta = 0, 1, \cdots\}$ implies the sequential completeness of $\{S(k), k = 0, 1, \cdots\}$.

To this end, the inequality (38) can be established by looking into the evolution of $\zeta_{i^*1}(k)$ for every $i^* \in \Omega_{\max}(k)$ and $\zeta_{i_*1}(k)$ for every $i_* \in \Omega_{\min}(k)$. This can be done using a similar analysis as in [29]. Details omitted due to space limitation. □

Remark 3.2: The kinematic model (1) was considered in the proposed formation control. With the aid of backstepping design [12], the result can be extended to deal with the distributed formation control for nonholonomic robots described by the following dynamic models [3]:

$$M(q)\ddot{q} + C(q,\dot{q})\dot{q} + G(q) = B(q)\tau + J^T(q)\lambda \quad (39)$$

$$J(q)\dot{q} = 0 \quad (40)$$

where $q = \begin{bmatrix} q_1 & \cdots & q_n \end{bmatrix}^T \in \Re^n$ is the generalized coordinates, $M(q) \in \Re^{n\times n}$ is a bounded positive-definite symmetric inertia matrix, $C(q,\dot{q}) \in \Re^{n\times n}$ is the centripetal and coriolis matrix, $G(q) \in \Re^n$ is the gravitation force vector, $B(q) \in \Re^{n\times r}$ is the input transformation matrix, $\tau \in \Re^r$ is the input vector of forces and torques, $J(q) \in \Re^{(n-m)\times n}$ is the matrix associated with the constraints, and $\lambda \in \Re^{n-m}$ is the vector of constraint forces on the contact point between the rigid body and the surface. ◇

IV. Simulation

In this section, we simulate the proposed formation control by considering three mobile robots moving according to a circular contour while maintaining a right triangle formation. We assume that the sensor/communication topologies are changing according to the sequence $\{S(t_k), k \in \aleph\}$ defined in (8), and robot 1 can receive the desired trajectory contour information $q_0(t)$.

Let $q_0(t)$ be $[\sin(0.2t), -\cos(0.2t)]^T$. The corresponding moving frame is given by $e_1(t) = [\cos(0.2t), -\sin(0.2t)]^T$, $e_2(t) = [\sin(0.2t), \cos(0.2t)]^T$. The formation parameters

are given by $\alpha_{11} = 0, \alpha_{12} = 0, \alpha_{21} = -1, \alpha_{22} = 1, \alpha_{31} = -1, \alpha_{32} = -1$. The initial conditions $[x_i(t_0), y_i(t_0), \theta_i(t_0)]$ are given by $[0.1, 0.2, \pi/4], [1, -2, \pi/6], [-1, -1.5, 0]^T$ for $i = 1, 2, 3$, respectively, $a_{i2}^k = 0.2$ and $T_s = 0.1$. Figure 2 illustrates the phase portrait under the proposed controls proposed controls (18) and (19).

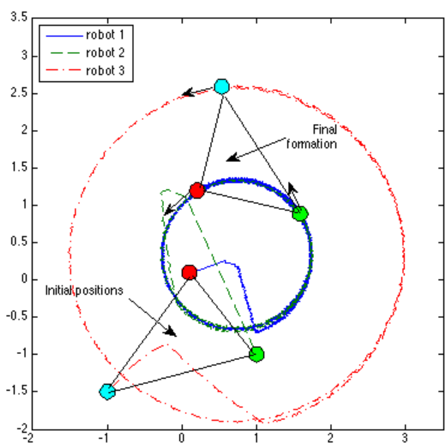

Fig. 2. Phase portrait of three robots

V. Conclusion

In this paper, we proposed a new nonlinear formation control method for nonholonomic mobile robots. Formation patterns are defined based on local coordinates with respect to the moving frame attached to the desired trajectory contour for the group. Through the design of a finite-time distributed observer for the desired trajectory, the formation control design can be done with limited information for the desired trajectory. System stability was rigorously proved using a contraction mapping method. Simulation result validated the effectiveness of the proposed design.

References

[1] A. Ailon and I. Zohar, "Control strategies for driving a group of nonholonomic kinematic mobile robots in formation along a time-parameterized path," *IEEE/ASME Transactions on Mechatronics*, vol. 17, pp. 326–336, 2012.

[2] M. Arcak, "Passivity as a design tool for group coordination," *IEEE Trans. Automat. Contr.*, vol. 52, pp. 1380–1389, 2007.

[3] A. Bloch and S. Drakunov, "Stabilization and tracking in the nonholonomic integrator via sliding modes," *System and contr. lett.*, vol. 29, pp. 91–99, 1996.

[4] Y. Cao, W. Ren, and Z. Meng, "Decentralized finite-time sliding mode estimators and their applications in decentralized finite-time formation tracking," *Systems & Control Letters*, vol. 59, pp. 522–529, 2010.

[5] J. Chen, D. Sun, J. Yang, and H. Chen, "Leader-follower formation control of multiple nonholonomic mobile robots incorporating a receding-horizon scheme," *The International Journal of Robotics Research*, vol. 29, pp. 727–747, 2010.

[6] D. V. Dimarogonal and K. J. Kyriakopoulos, "On the rendezvous problem for multiple nonholonomic agents," *IEEE Transactions on Automatic Control*, vol. 52, pp. 916–922, 2007.

[7] K. D. Do, "Formation tracking control of unicycle-type mobile robots with limited sensing range," *IEEE Trans. on Control System Technology*, vol. 16, pp. 527–538, 2008.

[8] J. A. Fax and R. M. Murray, "Information flow and cooperative control of vehicle formations," *IEEE Transactions on Automatic Control*, vol. 49, pp. 1465–1476, 2004.

[9] J. Ghommam, H. Mehrjerdi, M. Saad, and F. Mnif, "Formation path following control of unicycle-type mobile robots," *Robotics and Autonomous Systems*, vol. 58, pp. 722–736, 2010.

[10] A. Jadbabaie, J. Lin, and A. Morse, "Coordination of groups of mobile autonomous agents using nearest neighbor rules," *IEEE Trans. on Automatic Control*, vol. 48, pp. 988–1001, 2003.

[11] M. Ji and M. Egerstedt, "Distributed coordination control of multiagent systems while preserving connectedness," *IEEE Trans. on Robotics*, vol. 23, pp. 693–703, 2007.

[12] M. Krstic, I. Kanellakopoulos, and P. V. Kokotovic, *Nonlinear and Adaptive Control Design*. New York: Wiley, 1995.

[13] Z. Lin, M. Brouchke, and B. Francis, "Local control strategies for groups of mobile autonomous agents," *IEEE Trans. on Automatic Control*, vol. 49, pp. 622–629, 2004.

[14] Z. Lin, W. Ding, G. Yan, C. Yu, and A. Giua, "Leader-follower formation via complex laplacian," *Automatica*, vol. 49, p. 1900=1906, 2013.

[15] Z. Lin, B. Francis, and M. Maggiore, "Necessary and sufficient graphical conditions for formation control of unicycles," *IEEE Transactions on Automatic Control*, vol. 50, pp. 121–127, 2005.

[16] ——, "State agreement for continuous-time coupled nonlinear systems," *SIAM Journal on Control and Optimization*, vol. 46, pp. 288–307, 2007.

[17] S. Liu, D. Sun, and C. Zhu, "Coordinated motion planning for multiple mobile robots along designed paths with formation requirement," *IEEE/ASME Transactions on Mechatronics*, vol. 16, pp. 1021–1031, 2011.

[18] Q. Lu, S. Liu, X. Xie, and J. Wang, "Decision making and finite-time motion control for a group of robots," *IEEE Trans. on Cybernetics*, vol. 43, pp. 738–750, 2013.

[19] L. Moreau, "Stability of multiagent systems with time-dependent communication links," *IEEE Transactions on Automatic Control*, vol. 50, pp. 169–182, 2005.

[20] N. Moshtagh and A. Jadbabaie, "Distributed geodesic control laws for flocking of nonholonomic agents," *IEEE Transactions on Automatic Control*, vol. 52, pp. 681–686, 2007.

[21] R. M. Murray and S. S. Sastry, "Nonholonomic motion planning: Steering using sinusoids," *IEEE Trans. on Auto. Contr.*, vol. 38, pp. 700–716, 1993.

[22] B. S. Park, J. B. Park, and Y. H. Choi, "Adaptive formation control of electrically driven nonholonomic mobile robots with limited information," *IEEE Trans. on Systems, Man, and Cybernetics – Part B: Cybernetics*, vol. 41, pp. 1061–1075, 2011.

[23] Z. Qu, *Cooperative Control of Dynamical Systems*. London: Springer-Verlag, 2009.

[24] Z. Qu, J. Wang, and R. A. Hull, "Cooperative control of dynamical systems with application to autonomous vehicles," *IEEE Transactions on Automatic Control*, vol. 53, pp. 894–911, 2008.

[25] W. Ren and R. W. Beard, *Distributed Consensus in Multi-vehicle Cooperative Control*. London: Springer-Verlag, 2008.

[26] W. Ren and N. Sorensen, "Distributed coordination architecture for multi-robot formation control," *Robotics and Autonomous Systems*, 2007.

[27] R. O. Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, pp. 215–233, 2007.

[28] R. Sepulchre, A. Paley, and N. E. Leonard, "Stabilization of planar collective motion: all-to-all communication," *IEEE Transactions on Automatic Control*, vol. 52, pp. 916–922, 2007.

[29] J. Wang, M. Obeng, Z. Qu, T. Yang, G. Staskevich, and B. Abbe, "Discontinuous cooperative control for consensus of multiagent systems with switching topologies and time-delays," in *52th IEEE Conf. on Dec. and Ctrl*, Florence, Italy, Dec 2013, pp. 7333–7338.

[30] J. Wang, Z. Qu, and M. Obeng, "A distributed cooperative steering control with application to nonholonomic robots," in *49th IEEE Conf. on Dec. and Ctrl*, Atlanta, GA, Dec 2010, pp. 4571–4576.

LIST OF SYMBOLS, ABBREVIATIONS, AND ACRONYMS

HJB	Hamilton-Jacobi-Bellman
RBF	Radial Basis Function
MPI	Multiagent Policy Iteration
$\equiv$	identically equal
$\triangleq$	defined as
$<$ $(>)$	less (greater) than
$\leq$ $(\geq)$	less (greater) than or equal to
$\forall$	for all
$\in$	belongs to
$\rightarrow$	tends to
$\sum$	summation
$\bigcup$	union
$\|x\|$	the norm of a vector x
$\max$	maximum
$\min$	minimum
$\Re^n$	the $n-$dimensional Euclidean space
diag$[x_1, \cdots, x_n]$	a diagonal matrix with diagonal elements x_1 to x_n
$\dot{x}$	the first derivative of x with respect to time
$\ddot{x}$	the second derivative of x with respect to time
$A^T(x^T)$	the transpose of a matrix A (a vector x)
argmin	the argument of the minimum
$\mathcal{L}_2$	the space defined based on $2-$norm
$\mathcal{L}_\infty$	the space defined based on $\infty-$norm
sgn$(\cdot)$	the signum function